國立臺灣大學 數學所 碩士論文 共整合向量模型之平均數-變異數動態投資組合方法探討 詹孟諭 撰 民國九十九年十二月

國立臺灣大學理學院數學系

碩士論文

Department of Graduate Institute of Mathematics

College of Science

National Taiwan University

Master Thesis

共整合向量模型之

平均數-變異數動態投資組合方法探討

Dynamic Approaches to

Mean-Variance Portfolio Selection

in Cointegrated Vector Autoregressive Systems

詹孟諭

Meng-Yu Chan

指導教授：劉淑鶯 博士

Advisor:Shu-Ing Liu, Ph.D.

中華民國 99 年 12 月

December, 2010

謝辭

終於完成論文了，碩士生活即將畫上句點，面對未知的未來，心中充滿期待也有著些許的不安，我告訴自己，就像爬山一樣，一步一步向前走就對了，不要抬頭去看要走的路到底有多長、是否好走，而慢了向前走的步伐。

回頭看碩士生活，論文真的花了我很多時間，而本論文得以順利完成，首先感謝我的指導教授　劉淑鶯博士，在撰寫論文的過程中，不斷地給我指導與協助，想跟我的老師說，孟諭不會因為時間的流逝，就淡忘了您對我的幫助。此外，在論文口試之際，感謝我的口試委員　彭柏堅博士與鄭天澤博士對論文提供諸多意見與建議，使本論文能更加完善，想跟我的二位口試老師說，謝謝您們在寒流來時，還撥空來參加我的口試，口試的過程，令我收獲很多。

最後，更要感謝所有家人全力地支持，讓我好好地完成學業，謝謝詹鴻呈先生、傅雪琴女士、陳光雄先生、林淑貞女士、貽芬姐、怡妗姐、瓏仁、岳鈞、雯琇。

陳小圈，謝謝你總是全力地支持我去追逐我的夢想，讓我無後顧之憂地享受我的生活。

詹孟諭　謹誌于

台灣大學數學所金融數學組

西元 2011.2.11.

摘要

本文的資產價格數列模型，採用 PCB 共整合模型，將它整理成 VAR(1)的型式，式子中的參數，必須符合本文的公式，然後我們藉由這個模型，使用 Markowitz(1952)提出的平均數－變異數最佳化方法(Mean－Variance optimization approach)－固定風險、求報酬最大化的目標下，延伸作多期資產配置，再分別對一階段方法、二階段方法作討論。在有交易成本及截距項的情況下，比較這二種方法，同樣風險下的淨期望報酬。另外也比較，在特殊例子下，一階段方法與二階段方法，何種方法在同樣的風險下，能為我們帶來較大的淨期望報酬。

關鍵詞：PCB 共整合模型、Markowitz 平均數－變異數最佳化方法、多期資產配置、一階段方法、二階段方法

Abstract

This paper uses the PCB Cointegration Model to organize the sequence information of the price of financial commodity into the VAR(1) type. The parameters in the formula VAR(1) must meet the formulas in this paper. We extend Markowitz's mean-variance optimization approach published in 1952, which is to maximize the return under the fixed risk, to multi-stage asset allocation, and use this new model to discuss the one-stage method and the two-stage method. The paper then compares the net returns of the two methods when undertaking the same risk, under the condition of transaction cost and intercept. We will also examine the one-stage method and the two-stage method in the special cases to determine which one can bring the better net expected return under the same risk.

Key Words: PCB Cointegration Model, Markowitz Mean-Variance Optimization Approach, Multi-stage Asset Allocation, One-stage Method, Two-stage Method

Contents

Chapter 1 Introduction 1
1.1 Motivation and purpose 1
1.2 Framework 5

Chapter 2 The Cointegration model.. 8

Chapter 3 Research Method 10
3.1 Notations 10
3.2 Objective functions 11
3.3 Dynamic portfolio selection methods 13
3.3.1 One-stage method 13
3.3.2 Two-stage method 15
3.4 Special cases 20

Chapter 4. Numerical Illustrations 25
4.1 Unit root test and cointegration test 25
4.2 Parameter estimation 26
4.3 Transaction costs 27
4.4 Numerical results 28
4.4.1 Results of one-stage method 28
4.4.2 Results of two-stage method 29
4.5 Method comparisons 30

Chapter 5 Conclusions and Recommendations 33

Reference 36

Appendix 38

Contents of Appendix

Appendix 1. One-stage method 38

Appendix1.1 · $\Lambda \neq 0; \beta \neq 0$ 38

Appendix1.1.1 · Obtained the initial estimate λ_1^* by Taylor expansion 38

Appendix1.1.2 · The identical equation sorts out from the formula $Var\left\langle w_{CO}^T r \middle| x_0 \right\rangle = \sigma_O^2$ 39

Appendix1.2 · $\Lambda \neq 0; \beta = 0$ 39

Appendix1.3 · $\Lambda = 0; \beta \neq 0$ 40

Appendix1.3.1 · One-stage method of the maximum return method.......40

Appendix1.3.2 · One-stage method of the minimum variance method......... 42

Appendix1.4 · $\Lambda = 0; \beta = 0$ 42

Appendix2. Two-stage method 43

Appendix2.1 · $\Lambda \neq 0; \beta \neq 0$ 43

Appendix2.1.1 · Deriving the vector of expected return and the matrices of variance and expected transaction cost.............................. 43

Appendix2.1.2 · Obtained the initial estimate λ_2^* by Taylor expansion 47

Appendix2.1.3 · The identical equation sorts out from the formula $Var\left\langle V^{*T} Z \middle| x_0 \right\rangle = \sigma_0^2$........ 48

Appendix2.2 · $\Lambda \neq 0; \beta = 0$ 49

Appendix2.3 · $\Lambda = 0; \beta \neq 0$ 51

Appendix2.3.1 · Two-stage method of the maximum return method.......51

Appendix2.3.2 · Two-stage method of the minimum variance method.…….52

Appendix2.4 · $\Lambda = 0; \beta = 0$ 55

Contents of Figure

Figure 1. Framework……………………………………………………………......7

Contents of Table

Table 1. Comparison of expected net return……..………………………….................31

Table 2. Comparison of expected net return……..………………………….................32

Table 3. Comparison of expected net return……..………………………….................33

Table 4. Comparison of net return and the return per risk unit under $\Lambda \neq 0; \beta \neq 0$ ……..…………………………...33

1 Introduction

1.1 Motivation and purpose

Today, in our society, accumulation of wealth could be a difficult task for most people due to generally low salary structure, high commodity price, and low interest rate. Hence, investment turns out to be one of the effective ways to solve this issue. According to Brinson, Singer, and Beebower (1991), the main factors affecting investment returns are asset allocation, entry timing, and stock picking, among which asset allocation is the most influential one. In fact, the degree of influence of asset allocation to the return of investment can be up to 91.5%. For economists or investors who want to have satisfactory investment return figures, asset allocation is a very important topic.

The main concept of traditional asset allocation is derived from the Mean-Variance Portfolio Model published by Markowitz in 1952. This model deals mainly with the asset allocation problem of a single period. However, today, in our modern society, Markowitz's model has become insufficient to cope with the rapidly changing investment environment. In this paper, we first extend the Mean-Variance Portfolio Model to two-period and then combine with other models to address the shortcoming of the original model. Below is a brief introduction about the concept of the asset allocation of single and multi-period. Asset allocation of single period, also known as the traditional M-V model: No matter how the investment environment changes, investors make investment decision only at the beginning of the period. Asset allocation of multi-period: In several successive investment phases, investors will adjust the allocation strategy according to the latest information after each period. The new and much needed strategy is to readjust allocation weight in order to maximize the total return of their investments or to minimize the overall risk of the investments.

Early discussion about asset allocation involved using past data as a basis for future decision-making. However, Elam & Dixon (1988) pointed out that price series data of the financial commodity are often non-stationary. The price showed random walk, which means that price series of this period is not affected by that of the previous period, so using past data to predict future price is not efficient. In addition, as mentioned in the articles of Michaud (1989), Chopra & Ziemba (1993), Scherer (2002), and Wolf (2006), estimation error of the mean, variance and covariance will have a tremendous impact on the optimal portfolio. Therefore, information for future decision-making is considered very important.

If a variable appears as a random walk type, then it is considered a non-stationary variable. Granger and Newbold (1974) found that there would be spurious regression phenomenon among non-stationary variables. Traditionally in the empirical research, people mainly use regression methods to estimate the causal relationship between variables. However, overlooking spurious regression might lead to the misinterpretation of the result of empirical research. Hence, this fact that check whether variables are in steady state cannot be ignored when time series variables are used for the study. Therefore in this paper we check whether variables are in steady state first before we use time series variables for the empirical study.

Later studies found that many time series of economic or financial variables have non-stationary characteristic. Nelson and Plosser (1982) used a unit root test method to study macroeconomic data collected in the U.S. and found that most of the macroeconomic variables have the phenomenon of a unit root. In other words, most of the economic variables follow non-stationary process. In tradition, people would difference time series; however, important information might be lost as the result of differencing the time series. Granger (1983) proposed the concept of cointegration which uses cointegration vector to represent the relationship of the long-term mobile trend between the non-stationary time series. When the data appears to be cointegrated, it means that there is some information which may have a linear relationship. Thus the data doesn't need to go through the difference to be converted into stationary time series, which solves the problem of information loss from differencing the time series. Engle and Granger (1987) proposed the cointegration theory. According to their theory, if there is cointegration phenomenon in the regression relationship between non-stationary variables, then such a regression relationship still has economic significance. Therefore, this paper uses the PCB Cointegration Model to organize the sequence information of the price of financial commodity into the VAR (1) type. However, the parameters in the formula VAR(1) must meet the formulas(4)、(5)、(6) in this paper. Then we solve the problem of the future decision-making information with this model.

Traditionally, transaction cost is often overlooked in the discussion of asset allocation. When examining multi-period asset allocation, one should worry about the transaction cost becoming larger than the investment return due to frequent trading. In fact, transaction cost is always imposed when trading financial products. In order to model the common practice, transaction cost will be included in the following discussion.

This paper uses the above model to compare the one-stage method with the two-stage method: under the same risk, which method can bring a higher net return at the end of the two periods. Net return is defined as the total return subtracting transaction cost. With the one-stage method, the investment decision for both stages is made at the beginning of the first stage based on the information available at time. With the two-stage method, we make the investment decision for the first stage at the beginning of the stage according to the information available at time; and when the first phase is ended, we then use the new information available to make the best decision for the second phase.

1.2 Framework

This article is divided into five chapters. A brief description of the content in each chapter is as follows:

Chapter 1 Introduction

This chapter describes the motivation, purpose, and structure of the research in this paper.

Chapter 2 The Cointegration Model

This chapter describes the cointegration model.

Chapter 3 Research Method

This paper uses formula(3) transformed from PCB Cointegration Model to organize the sequence information of the price of the financial commodity, and then uses the mean-variance optimization method to set the objective function for this paper. We examine the one-stage method and the two-stage method of the dynamic portfolio selection approach and determine which one can produce better net expected return when undertaking the same risk.

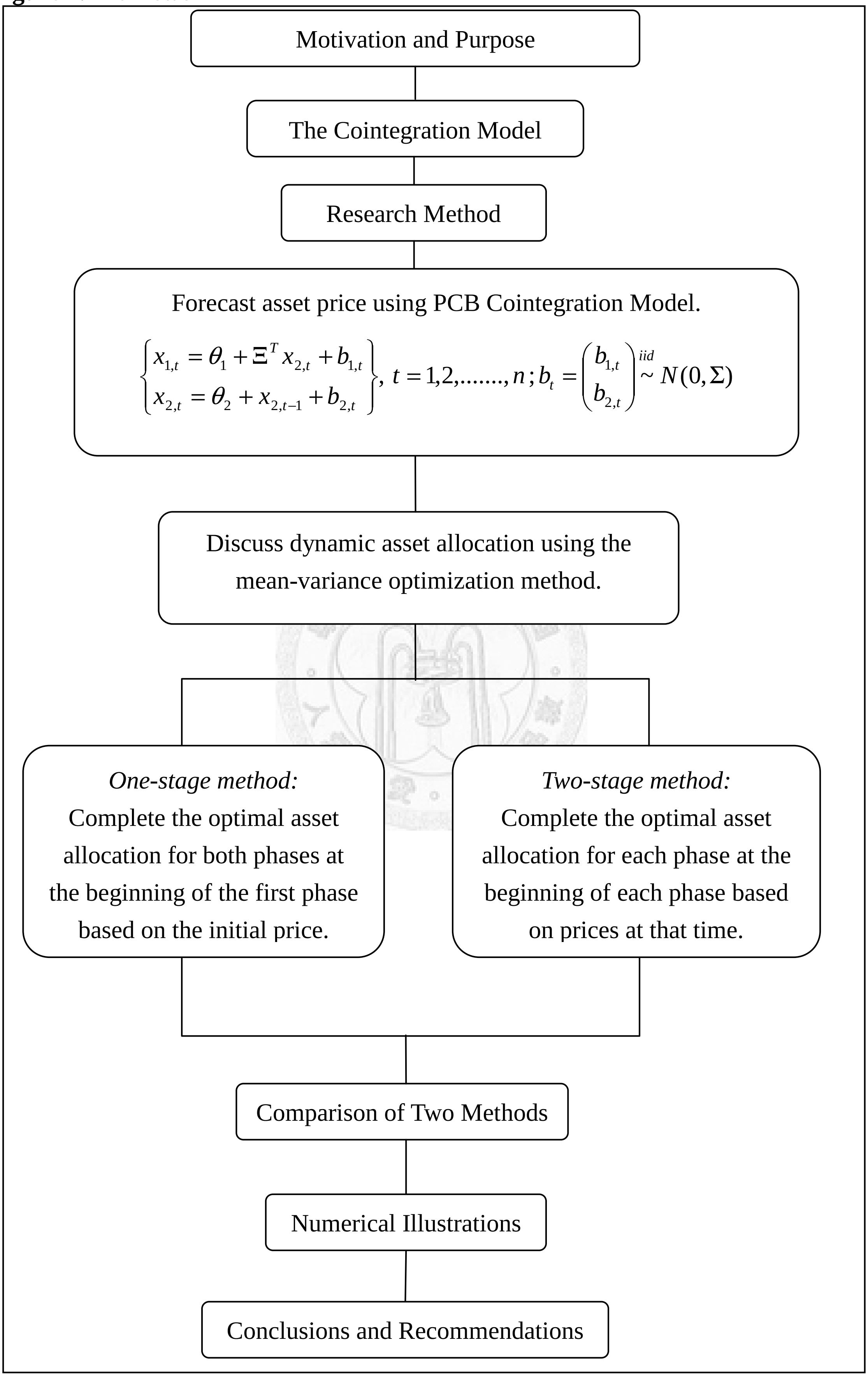
Chapter 4 Numerical Illustrations

Among the 19 sector indices in the Taiwan market, we pick 5 that are I(1) sequences then find their cointergration relationship. Next, estimate the model parameters and use numerical methods to compare one-stage method and two-stage method.

Chapter 5 Conclusions and Recommendations

We draw conclusions from the empirical results in chapter 4 and make recommendations of future research on this topic.

Figure 1. Framework



2 The Cointegration Model

In this chapter, we will show Cointegration Model which will be used to organize the sequence information of the price of the financial commodity in this paper.

Before being engaged in each type of statistical inference of the time series, we should first examine whether the sequence is in stationary state. When the time series is stationary, its data follows a stochastic process, but the probability distribution of this stochastic process does not change along with time; otherwise, this time series is called non-stationary time series.

Elam & Dixon (1988) pointed out that the price series of financial products usually show the characteristic of non-stationary process. That means the price sequence is a random walk type in which the price sequence of current period is not affected by the price sequence of the previous period. If a time series is non-stationary, the commonly used method to convert it to stationary is to take the difference on the variable or eliminate the tendency item. However, the conversion could potentially eliminate the implicit long-term message from the data, so to avoid this potential problem, we could adopt the cointegration concept proposed by Granger. According to the cointegration concept, a stable long-term balanced relation might exist between unstable variables, and this relation might cause synchronized tendency on the variables. Therefore, we use formula(3) transformed from PCB Cointegration Model to model price sequence of finance commodity in this paper.

PCB Cointegration model

$$\begin{Bmatrix} x_{1,t} = \theta_1 + \Xi^T x_{2,t} + b_{1,t} \\ x_{2,t} = \theta_2 + x_{2,t-1} + b_{2,t} \end{Bmatrix}, t = 1,2,......,n; \ \ b_t = \begin{pmatrix} b_{1,t} \\ b_{2,t} \end{pmatrix} \overset{iid}{\sim} N(0,\Sigma) \tag{1}$$

First, move $x_{1,t}$ 、 $x_{2,t}$ to the left of formula(1) as below:

$$\begin{pmatrix} x_{1,t} - \Xi^T x_{2,t} = \theta_1 + b_{1,t} \\ x_{2,t} = \theta_2 + x_{2,t-1} + b_{2,t} \end{pmatrix} \tag{2}$$

Next, rewrite Eq.(2) in vector and matrix forms.

$$\begin{pmatrix} I & -\Xi^T \\ 0 & I \end{pmatrix}\begin{pmatrix} x_{1,t} \\ x_{2,t} \end{pmatrix} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} x_{1,t-1} \\ x_{2,t-1} \end{pmatrix} + \begin{pmatrix} b_{1,t} \\ b_{2,t} \end{pmatrix}$$

$$\begin{pmatrix} x_{1,t} \\ x_{2,t} \end{pmatrix} = \begin{pmatrix} I & -\Xi^T \\ 0 & I \end{pmatrix}^{-1}\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \begin{pmatrix} I & -\Xi^T \\ 0 & I \end{pmatrix}^{-1}\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} x_{1,t-1} \\ x_{2,t-1} \end{pmatrix} + \begin{pmatrix} I & -\Xi^T \\ 0 & I \end{pmatrix}^{-1}\begin{pmatrix} b_{1,t} \\ b_{2,t} \end{pmatrix}$$

$$= \begin{pmatrix} I & \Xi^T \\ 0 & I \end{pmatrix}\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \begin{pmatrix} I & \Xi^T \\ 0 & I \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} x_{1,t-1} \\ x_{2,t-1} \end{pmatrix} + \begin{pmatrix} I & \Xi^T \\ 0 & I \end{pmatrix}\begin{pmatrix} b_{1,t} \\ b_{2,t} \end{pmatrix}$$

$$= \beta + \begin{pmatrix} 0 & \Xi^T \\ 0 & I \end{pmatrix}\begin{pmatrix} x_{1,t-1} \\ x_{2,t-1} \end{pmatrix} + u_t$$

$$= \beta + \Pi \begin{pmatrix} x_{1,t-1} \\ x_{2,t-1} \end{pmatrix} + u_t \tag{3}$$

$x_t = \beta + \Pi x_{t-1} + u_t$, its form is similar to VAR(1). Its coefficient vector and matrix, $\beta \,\&\, \Pi$, and random vector, u_t , are as follows:

$$\beta = \begin{pmatrix} I & \Xi^T \\ 0 & I \end{pmatrix}\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \tag{4}$$

$$\Pi = \begin{pmatrix} 0 & \Xi^T \\ 0 & I \end{pmatrix} \quad \text{where the property of } \Pi:$$

$$\Pi_1 = \Pi - I; \quad \Pi_1 = \begin{bmatrix} -I & \Xi^T \\ 0 & 0 \end{bmatrix}; \quad \Pi^k = \Pi, \quad k = 1,2,3..........; \quad \Pi_1\Pi = 0 \tag{5}$$

$$u_t = \begin{pmatrix} I & \Xi^T \\ 0 & I \end{pmatrix}\begin{pmatrix} b_{1,t} \\ b_{2,t} \end{pmatrix} \overset{iid}{\sim} N(0,\Sigma^*) \quad \text{where } \Sigma^* = \begin{pmatrix} I & \Xi^T \\ 0 & I \end{pmatrix}\Sigma\begin{pmatrix} I & \Xi^T \\ 0 & I \end{pmatrix}^T \tag{6}$$

3 Research Method

3.1 Notations

x_t denotes the random vector of asset prices taken logarithm at time t. According to formula(3), $x_t = \beta + \Pi x_{t-1} + u_t$. When $t \geq 2$, the x_t can be rewritten as:

$$x_t = \beta + (t-1)\Pi\beta + \Pi x_0 + \Pi\sum_{i=1}^{t-1} u_i + u_t \tag{7}$$

r_t , $r_t = x_t - x_{t-1}$, denotes the random vector of asset return during the time interval, $(t-1, t)$ .According to formula(3):

$$r_t = x_t - x_{t-1} = \beta + \Pi x_{t-1} + u_t - x_{t-1} = \beta + \Pi_1 x_{t-1} + u_t \tag{8}$$

When $t \geq 2$, the x_t can be rewritten as:

$$r_t = x_t - x_{t-1} = \Pi\beta + \Pi_1 u_{t-1} + u_t \tag{9}$$

State : Each x_t represents state t.

Stage : Stage t represents the range between x_{t-1} and x_t. Stage is also called period.

w_t denotes the weight of asset allocation between x_{t-1} and x_t. This is the weight that we allocate at stage t. When we know the price of assets x_t, we should know all the weights that are less than or equal to that at stage $t+1$. For example, when we know the price of assets x_1, we should know the price of assets x_0, then we can know the weights w_1 and w_2. In addition, positive weight represents buying of stocks; negative weight means selling or short selling of stocks. Prices of financial products do not always go up; they may go down as well. Short selling is the trading option to maximize profits during down time, thus we allow short selling in this paper. Because our assumption allows short selling of stocks, the sum of the weights in each stage does not need to be equal to one.

Λ denotes the fees charged for different trading methods, such as buying, selling, and short selling of stocks. It is determined based on the weight. For the convenience of problem solving and calculation, we represent it using a fixed ratio. And we assume that Λ is a symmetric positive definite matrix. The model of the transaction cost in this paper is $\frac{1}{2}w^T \Lambda w$. The reason we use quadratic form to present the model of transaction costs in this paper is to avoid the transaction costs becoming negative. Because we allow negative weight, we must use the quadratic form to ensure that transaction cost is positive.

Since we apply two-period dynamic asset allocation, the maximum number of $t = 2$.

3.2 Objective functions

Markowitz's mean-variance portfolio optimization methods provide two approaches: (1) under fixed risk, find the maximum return or (2) under fixed remuneration, minimize the risk. The method that we chose in this paper is the former, which is to find maximum return under fixed risk. In the special case of the following section, we will verify that both methods give the same return when undertaking a unit risk.

This paper uses two-stage mean-variance portfolio optimization framework, and the objective function is as follows:

$$f_C(w_1, w_2) = \max_{w_1, w_2} E\left\langle w_1^T r_1 + w_2^T r_2 \,\middle|\, x_0 \right\rangle - \frac{1}{2} E\left\langle w_1^T \Lambda w_1 + (w_2 - w_1)^T \Lambda (w_2 - w_1) \,\middle|\, x_0 \right\rangle \quad (10)$$

Restriction is as follow:

$$Var\left\langle w_1^T r_1 + w_2^T r_2 \,\middle|\, x_0 \right\rangle = \sigma_0^2$$

That means we want to optimize the net expected return of two-period, which is deducting transaction cost under the constraint of restricting the variance of total return to σ_0^2 at the end.

The objective function (10) and its restriction in this case which use the Lagrange multipliers method can be rewritten as:

$$f_C(w_1, w_2) = \max_{w_1, w_2} E\left\langle w_1^T r_1 + w_2^T r_2 \,\middle|\, x_0 \right\rangle - \frac{1}{2} E\left\langle w_1^T \Lambda w_1 + (w_2 - w_1)^T \Lambda (w_2 - w_1) \,\middle|\, x_0 \right\rangle$$
$$- \frac{\lambda_C}{2}\left[Var\left\langle w_1^T r_1 + w_2^T r_2 \,\middle|\, x_0 \right\rangle - \sigma_0^2\right] \quad (11)$$

where λ_C is a Lagrange multiplier.

Below we will discuss one-stage method and two-stage method respectively.

3.3 Dynamic portfolio selection methods

3.3.1 One-stage method

At state x_0, we optimize asset allocation for two periods according to asset prices at state x_0.

First, the objective function sort out from (11) is as follow:

$$\begin{aligned} f_{CO}(w) &= E\left\langle w^T r \,\middle|\, x_0 \right\rangle - \frac{1}{2} w^T \Lambda^* w - \frac{\lambda_1}{2}\left[Var\left\langle w^T r \,\middle|\, x_0 \right\rangle - \sigma_0^2\right] \\ &= w^T E\left\langle r \,\middle|\, x_0 \right\rangle - \frac{1}{2} w^T \Lambda^* w - \frac{\lambda_1}{2}\left[w^T Var\left\langle r \,\middle|\, x_0 \right\rangle w - \sigma_0^2\right] \\ &= w^T A - \frac{1}{2} w^T \Lambda^* w - \frac{\lambda_1}{2}\left[w^T B w - \sigma_0^2\right] \end{aligned} \quad (12)$$

The vectors and matrices represent respectively:

$$w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} ; \quad r = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} ; \quad \Lambda^* = \begin{bmatrix} 2\Lambda & -\Lambda \\ -\Lambda & \Lambda \end{bmatrix} ;$$

$$A = E\left\langle r \,\middle|\, x_0 \right\rangle = \begin{bmatrix} E\left\langle r_1 \,\middle|\, x_0 \right\rangle \\ E\left\langle r_2 \,\middle|\, x_0 \right\rangle \end{bmatrix} ;$$

$$B = Var\left\langle r \,\middle|\, x_0 \right\rangle = \begin{bmatrix} Var\left\langle r_1 \,\middle|\, x_0 \right\rangle & Cov\left\langle r_1, r_2 \,\middle|\, x_0 \right\rangle \\ Cov\left\langle r_2, r_1 \,\middle|\, x_0 \right\rangle & Var\left\langle r_2 \,\middle|\, x_0 \right\rangle \end{bmatrix}$$

Then to optimize the objective function, we perform the first order differential equal to zero on it.

$$A - \Lambda^* w - \lambda_1 B w = 0$$

$$w_{CO} = \Psi^{-1} A \qquad \text{where} \quad \Psi = \left(\Lambda^* + \lambda_1 B\right) \quad (13)$$

According to 3.1 (8)、(9), the results are derived as follows:

$$E\langle r_1 | x_0 \rangle = E\langle \beta + \Pi_1 x_0 + u_1 | x_0 \rangle = \beta + \Pi_1 x_0$$

$$E\langle r_2 | x_0 \rangle = E\langle \Pi\beta + \Pi_1 u_1 + u_2 | x_0 \rangle = \Pi\beta$$

$$Var\langle r_1 | x_0 \rangle = Var\langle \beta + \Pi_1 x_0 + u_1 | x_0 \rangle = \Sigma^*$$

$$Var\langle r_2 | x_0 \rangle = Var\langle \Pi\beta + \Pi_1 u_1 + u_2 | x_0 \rangle = D \quad \text{where} \quad D = \Pi_1 \Sigma^* \Pi_1^T + \Sigma^*$$

$$Cov\langle r_1, r_2 | x_0 \rangle = Cov\langle \beta + \Pi_1 x_0 + u_1, \Pi\beta + \Pi_1 u_1 + u_2 | x_0 \rangle$$

$$= Cov\langle u_1, \Pi_1 u_1 | x_0 \rangle = \Sigma^* \Pi_1^T$$

Since weight w_{CO} (13) contains an unknown figure λ_1, we try to use the equation of constraint to solve λ_1. In fact, in the process of finding solution, we will see λ_1 in the anti-function, and there is no way to find solution from that function directly. Therefore, our approach is to get a starting estimate λ_1^* by using Taylor expansion first. (The detailed derivations are in Appendix1.1.1.)

$$\lambda_1^* \approx \frac{A^T \cdot \Lambda^{*^{-1}} \cdot B \cdot \Lambda^{*^{-1}} \cdot A - \sigma_0^2}{2A^T \cdot \Lambda^{*^{-1}} \cdot B \cdot \Lambda^{*^{-1}} \cdot B \cdot \Lambda^{*^{-1}} \cdot A} \tag{14}$$

Then from the formula $Var\left\langle w_{CO}^T r \middle| x_0 \right\rangle = \sigma_O^2$, we can sort out an identity equation. (The detailed derivations are in Appendix1.1.2.)

$$A^T \cdot \Psi^{-1} \cdot A - A^T \cdot \Psi^{-1} \cdot \Lambda^* \cdot \Psi^{-1} \cdot A = \lambda_1 \cdot \sigma_0^2 \tag{15}$$

Next, we take the initial estimated value λ_1^* into the function Ψ^{-1} of the left side of Eq. (15), and then we can get a new estimated value from the right side of the equation. We then confirm if the difference between the initial estimated value and the new estimated value is less than the estimated error of our setting. If not, we take the new estimated value and substitute it into the left side of the formula until the values of two sides are very close, expressing convergence, then we stop. If so, we can get an answer directly, which results in a more accurate estimated value λ_1^S.

The optimal weight of one stage method:

$$w_{CO}^{*} = \Psi^{*^{-1}} A \qquad \text{where} \quad \Psi^{*} = \left(\Lambda^{*} + \lambda_1^S B\right) \tag{16}$$

Expected return after deducting transaction cost of one-stage method as follows:

$$\begin{aligned} E\left\langle {w_{CO}^{*}}^{T} r \middle| x_0 \right\rangle - \frac{1}{2} {w_{CO}^{*}}^{T} \Lambda^{*} w_{CO}^{*} &= {w_{CO}^{*}}^{T} A - \frac{1}{2} {w_{CO}^{*}}^{T} \Lambda^{*} w_{CO}^{*} \\ &= A^T \cdot \Psi^{*^{-1}} \cdot A - \frac{1}{2} A^T \cdot \Psi^{*^{-1}} \cdot \Lambda^{*} \cdot \Psi^{*^{-1}} \cdot A \\ &= A^T \cdot \Psi^{*^{-1}} \cdot A - \frac{1}{2}\left(A^T \cdot \Psi^{*^{-1}} \cdot A - \lambda_1^S \cdot \sigma_0^2\right) \\ &= \frac{1}{2}\left(A^T \cdot \Psi^{*^{-1}} \cdot A + \lambda_1^S \cdot \sigma_0^2\right) \end{aligned}$$

3.3.2 Two-stage method

At beginning of the first period, we will allocate the weight of asset based on the initial price. We will then re-allocate the weight of asset again based on the asset price at beginning of the second period.

In this method, we use the backward way to solve it. In other words, we get the optimal weight of the second phase by using the objective function of the second phase.

Then we apply the optimal weight of the second stage to the overall objective function to get the optimal weight of the first phase.

In this paper, we don't consider the transaction cost in the objective function of the second stage. Instead, we use an undetermined coefficient a_1 to adjust the change.

Step1.

Suppose x_0 and x_1 are known, the objective function of the second phase is as follow:

$$f_2(w_2) = E\langle w_2^T r_2 \,|\, x_0, x_1 \rangle - \frac{1}{2} Var\langle w_2^T r_2 \,|\, x_0, x_1 \rangle$$

$$= w_2^T E\langle r_2 \,|\, x_0, x_1 \rangle - \frac{1}{2} w_2^T Var\langle r_2 \,|\, x_0, x_1 \rangle w_2$$

We want to obtain the optimal decision of the second stage, so we perform the first order differential equal to zero on it.

$$E\langle r_2 \,|\, x_0, x_1 \rangle - Var\langle r_2 \,|\, x_0, x_1 \rangle w_2 = 0$$

$$w_2^* = \left(Var\langle r_2 \,|\, x_0, x_1 \rangle\right)^{-1} E\langle r_2 \,|\, x_0, x_1 \rangle$$

According to 3.1 (8)、(9), the results are derived as follows:

$$E\langle \mathrm{r}_2 \,|\, x_0, x_1 \rangle = E\langle \beta + \Pi_1 x_1 + u_2 \,|\, x_0, x_1 \rangle = \beta + \Pi_1 x_1$$

$$Var\langle \mathrm{r}_2 \,|\, x_0, x_1 \rangle = Var\langle \beta + \Pi_1 x_1 + u_2 \,|\, x_0, x_1 \rangle = \Sigma^*$$

The deduced result is obtained:

$$w_2^* = \gamma + \Phi x_1 \qquad \text{where} \quad \gamma = \left(\Sigma^*\right)^{-1} \beta \;\; ; \;\; \Phi = \left(\Sigma^*\right)^{-1} \Pi_1$$

We apply factor a_1 then it becomes:

$$a_1 w_2^* = a_1(\gamma + \Phi x_1) \tag{17}$$

Step2.

Next, we must take the optimal weight of the second phase (17), which multiplies an undetermined coefficient a_1, to substitute it into the optimal objective function of two periods (11), as shown below:

$$f_{CT}(w_1, a_1) = E\left\langle w_1^T r_1 + a_1 w_2^{*T} r_2 \middle| x_0 \right\rangle - \frac{1}{2} E\left\langle w_1^T \Lambda w_1 + \left(a_1 w_2^* - w_1\right)^T \Lambda \left(a_1 w_2^* - w_1\right) \middle| x_0 \right\rangle$$
$$- \frac{\lambda_2}{2}\left[Var\left\langle w_1^T r_1 + a_1 w_2^{*T} r_2 \middle| x_0 \right\rangle - \sigma_0^2 \right]$$

The objective function is sorted out as following:

$$\begin{aligned} f_{CT}(V) &= E\left\langle V^T Z \middle| x_0 \right\rangle - \frac{1}{2} E\left\langle V^T C V \middle| x_0 \right\rangle - \frac{\lambda_2}{2}\left[Var\left\langle V^T Z \middle| x_0 \right\rangle - \sigma_0^2 \right] \\ &= V^T E\left\langle Z \middle| x_0 \right\rangle - \frac{1}{2} V^T E\langle C | x_0 \rangle V - \frac{\lambda_2}{2}\left[V^T Var\langle Z | x_0 \rangle V - \sigma_0^2 \right] \\ &= V^T F - \frac{1}{2} V^T H V - \frac{\lambda_2}{2}\left[V^T G V - \sigma_0^2 \right] \end{aligned} \tag{18}$$

The vectors and matrices represent respectively:

$$V = \begin{bmatrix} w_1 \\ a_1 \end{bmatrix} ; \quad Z = \begin{bmatrix} r_1 \\ w_2^{*T} r_2 \end{bmatrix} ; \quad C = \begin{bmatrix} 2\Lambda & -\Lambda w_2^* \\ -w_2^{*T}\Lambda & w_2^{*T}\Lambda w_2^* \end{bmatrix} ;$$

$$F = E\left\langle Z \middle| x_0 \right\rangle = \begin{bmatrix} E\langle r_1 | x_0 \rangle \\ E\left\langle w_2^{*T} r_2 \middle| x_0 \right\rangle \end{bmatrix} ;$$

$$H = E\left\langle C \middle| x_0 \right\rangle = \begin{bmatrix} E\langle 2\Lambda | x_0 \rangle & -E\left\langle \Lambda w_2^* \middle| x_0 \right\rangle \\ -E\left\langle w_2^{*T}\Lambda \middle| x_0 \right\rangle & E\left\langle w_2^{*T}\Lambda w_2^* \middle| x_0 \right\rangle \end{bmatrix} ;$$

$$G = Var\left\langle Z \middle| x_0 \right\rangle = \begin{bmatrix} Var\langle r_1 | x_0 \rangle & Cov\left\langle r_1, w_2^{*T} r_2 \middle| x_0 \right\rangle \\ Cov\left\langle w_2^{*T} r_2, r_1 \middle| x_0 \right\rangle & Var\left\langle w_2^{*T} r_2 \middle| x_0 \right\rangle \end{bmatrix}$$

After the first order differential, we get:

$$F - HV - \lambda_2 G V = 0$$

$$V^* = \Omega^{-1}F \quad \text{where} \quad \Omega = \left(H + \lambda_2 G\right) \tag{19}$$

According to 3.1 (8)、(9), the results are derived as follows:

(The detailed derivations are in Appendix2.1.1.)

$$E\langle r_1 | x_0 \rangle = E(\beta + \Pi_1 x_0 + u_1) = \beta + \Pi_1 x_0$$

$$E\left\langle w_2^{*T} r_2 \middle| x_0 \right\rangle = \ell \quad where \quad \ell = L^T \Pi \beta + tr(\varphi); \;\; L = (\gamma + \Phi\beta); \;\; \varphi = \Phi^{\mathrm{T}} \Pi_1 \Sigma^*$$

$$Var\langle r_1 | x_0 \rangle = Var\langle \beta + \Pi_1 x_0 + u_1 | x_0 \rangle = \Sigma^*$$

$$Var\left\langle w_2^{*T} r_2 \middle| x_0 \right\rangle = k \quad where \quad D = \Pi_1 \Sigma^* \Pi_1^T + \Sigma^* \;; \quad \varphi = \Phi^{\mathrm{T}} \Pi_1 \Sigma^*$$

$$k = \gamma^T D \gamma + \left(\gamma^T + L^{\mathrm{T}}\right) D \Phi \beta + 2tr\left(\varphi^2\right) + tr(\varphi) + \left(\beta^T \Pi^T \Phi + 2L^T \Pi_1\right) \Sigma^* \Phi^{\mathrm{T}} \Pi \beta$$

$$Cov\left\langle r_1, w_2^{*T} r_2 \middle| x_0 \right\rangle = N \quad where \quad N = \Sigma^* \left(\Pi_1^T \gamma + \Pi_1^T \Phi \beta + \Phi^{\mathrm{T}} \Pi \beta\right)$$

$$E\langle 2\Lambda | x_0 \rangle = 2\Lambda$$

$$E\left\langle \Lambda w_2^* \middle| x_0 \right\rangle = \Lambda L \quad where \quad L = (\gamma + \Phi\beta)$$

$$E\left\langle w_2^{*T} \Lambda w_2^* \middle| x_0 \right\rangle = L^T \Lambda L + tr(\varphi_1) \quad where \quad L = (\gamma + \Phi\beta) \;; \quad \varphi_1 = \Phi^{\mathrm{T}} \Lambda \Phi \Sigma^*$$

Same as one-stage method, the vector V^* of two-stage method also contains unknown figure λ_2 and there is no way to utilize equation of constraint to solve λ_2. Hence, we also get a starting estimate λ_2^* by Taylor expansion.

(The detailed derivations are in Appendix2.1.2.)

$$\lambda_2^* \approx \frac{F^T \cdot H^{-1} \cdot G \cdot H^{-1} \cdot F - \sigma_0^2}{2F^T \cdot H^{-1} \cdot G \cdot H^{-1} \cdot G \cdot H^{-1} \cdot F} \tag{20}$$

Then from the formula $Var\left\langle V^{*T} Z \middle| x_0 \right\rangle = \sigma_0^2$, we can sort out an identity equation.

(The detailed derivations are in Appendix2.1.3.)

$$F^T \cdot \Omega^{-1} \cdot F - F^T \cdot \Omega^{-1} \cdot H \cdot \Omega^{-1} \cdot F = \lambda_2 \cdot \sigma_0^2 \tag{21}$$

Next, we take the initial estimated value λ_2^* into the function Ω^{-1} of the left side of Eq.(21), so that we can get a new estimated value of the right side of the equation. We then confirm if the difference of the initial estimated value λ_2^* and the new estimated value is less than the estimated error of our setting. If not, we take the new estimated value and substitute it into the left side of the formula until the value of two sides are very close, expressing convergence, then we stop; If so, we can get an answer directly, which would result in a more accurate estimated value λ_2^S.

The two-stage method obtains a vector V_λ^* which is combination of the optimal weight of the first stage and a scale factor. $V_\lambda^* = \Omega^{*-1} F$ where $\Omega^* = \left(H + \lambda_2^S G\right)$ (22)

Expected return of two-stage method minus transition cost is represented as follows:

$$
\begin{aligned}
& E\left\langle V_\lambda^{*T} Z \,\middle|\, x_0 \right\rangle - \frac{1}{2} E\left\langle V_\lambda^{*T} C V_\lambda^* \,\middle|\, x_0 \right\rangle \\
&= V_\lambda^{*T} F - \frac{1}{2} V_\lambda^{*T} H V_\lambda^* \\
&= F^T \cdot \Omega^{*-1} \cdot F - \frac{1}{2} F^T \cdot \Omega^{*-1} \cdot H \cdot \Omega^{*-1} \cdot F \\
&= F^T \cdot \Omega^{*-1} \cdot F - \frac{1}{2}\left(F^T \cdot \Omega^{*-1} \cdot H - \lambda_2^S \cdot \sigma_0^2\right) \\
&= \frac{1}{2}\left(F^T \cdot \Omega^{*-1} \cdot F + \lambda_2^S \cdot \sigma_0^2\right)
\end{aligned}
$$

3.4 Special Cases

In this section, we will show four special cases.

Case #1: The first example is the typical Markowitz mean-variance portfolio optimization approach in which asset allocation decisions are made at the beginning of the investment period according to the price at the time regardless of the length of the investment period. In other words, the asset allocation weight will not be changed through time. This method is also known as static method. In this paper, we will transform the PCB Cointegration Model to formula(3), and then incorporate the new model together with the transaction cost model into the static method.

The objective function of the static method is as follows:

$$
\begin{aligned}
f_{\mathrm{C}}(w) &= \max_{w} E\left\langle w^T r \,\middle|\, x_0 \right\rangle - \frac{1}{2} w^T \Lambda w - \frac{\lambda}{2}\left[Var\left\langle w^T r \,\middle|\, x_0 \right\rangle - \sigma_0^2\right] \\
&= w^T A_S - \frac{1}{2} w^T \Lambda w - \frac{\lambda}{2}\left[w^T B_S w - \sigma_0^2\right]
\end{aligned}
$$

According to the definition of return in 3.1 (8)、(9), the return random vector is as follows:

$$
\begin{aligned}
r = x_2 - x_0 &= (x_2 - x_1) + (x_1 - x_0) = r_2 + r_1 = \Pi\beta + \Pi_1 u_1 + u_2 + \beta + \Pi_1 x_0 + u_1 \\
&= \Pi_1 x_0 + (\Pi + I)\beta + (\Pi_1 + I)u_1 + u_2 = \Pi_1 x_0 + (\Pi + I)\beta + \Pi u_1 + u_2 \qquad (23)
\end{aligned}
$$

After taking the first order differential, we get:

$$
w_S = \mathrm{M}^{-1} A_S \qquad \text{where} \quad \mathrm{M} = (\Lambda + \lambda B_S) \qquad (24)
$$

Result derived according to Eq.(23) is as follows:

$$E\langle r \mid x_0 \rangle = E\langle \Pi_1 x_0 + (\Pi + I)\beta + \Pi u_1 + u_2 \mid x_0 \rangle = \Pi_1 x_0 + (\Pi + I)\beta$$

$$Var\langle r \mid x_0 \rangle = Var\langle \Pi_1 x_0 + (\Pi + I)\beta + \Pi u_1 + u_2 \mid x_0 \rangle = \Pi\Sigma^*\Pi^{\mathrm{T}} + \Sigma^*$$

Then, with the same approach used previously to estimate value λ, we get a more accurate estimate λ_S, thus the expected return after deducting transaction cost of static method is as follows:

$$E\left\langle {w_S^*}^T r \middle| x_0 \right\rangle - \frac{1}{2} {w_S^*}^T \Lambda w_S^* = \frac{1}{2}\left(A_S^T \cdot \mathbf{M}^{*-1} \cdot A_S + \lambda_S \cdot \sigma_0^2\right)$$

Case #2: Given the condition $\Lambda \neq 0; \beta = 0$, compare the two dynamic methods to determine which method gives higher expected return after deducting transaction cost. (See Appendix 1.2、2.2 for the detailed derivations)

The objective function is as follows:

$$f_C(w_1, w_2) = \max_{w_1, w_2} E\left\langle w_1^T r_1 + w_2^T r_2 \mid x_0 \right\rangle - \frac{1}{2} E\left\langle w_1^T \Lambda w_1 + (w_2 - w_1)^T \Lambda (w_2 - w_1) \middle| x_0 \right\rangle - \frac{\lambda_{\mathrm{C}}}{2}\left[Var\left\langle w_1^T r_1 + w_2^T r_2 \mid x_0 \right\rangle - \sigma_0^2\right]$$

Expected return after deducting transaction cost of one-stage method is as follows:

$$E\left\langle {w_{CO}^*}^T r \middle| x_0 \right\rangle - \frac{1}{2} {w_{CO}^*}^T \Lambda^* w_{CO}^* = \frac{1}{2}\left(A_0^T \cdot \Psi^{*-1} \cdot A_0 + \lambda_3^S \cdot \sigma_0^2\right)$$

Expected return after deducting transaction cost of two-stage method is as follows:

$$E\left\langle {V_\lambda^*}^T Z \middle| x_0 \right\rangle - \frac{1}{2} E\left\langle {V_\lambda^*}^T C V_\lambda^* \middle| x_0 \right\rangle = \frac{1}{2}\left(F_0^T \cdot \Omega_0^{*-1} \cdot F_0^T + \lambda_4^S \cdot \sigma_0^2\right)$$

Case #3: Given the condition $\Lambda = 0; \beta \neq 0$, compare the two dynamic methods to determine which method gives higher expected return after deducting transaction cost. (See Appendix 1.3.1、2.3.1 for the detailed derivations)

The objective function is as follows:

$$f(w_1, w_2) = \max_{w_1, w_2} E\left\langle w_1^T r_1 + w_2^T r_2 \middle| x_0 \right\rangle - \frac{\lambda}{2}\left[Var\left\langle w_1^T r_1 + w_2^T r_2 \middle| x_0 \right\rangle - \sigma_0^2\right]$$

Expected return of one-stage method is as follows:

$$E\left\langle w_O^{*T} r \middle| x_0 \right\rangle = \sigma_O \sqrt{A^T B^{-1} A}$$

Expected return of two-stage method is as follows:

$$E\left\langle V_\lambda^{*T} Z \middle| x_0 \right\rangle = \sigma_O \sqrt{F^T G^{-1} F}$$

We will have further discussion on this case. In previous discussion, we talk about seeking maximum return while undertaking the fixed risk. We want to verify if the return per unit of risk remains the same when seeking for a minimum risk for a fixed return. We decided to use this case for discussion because the λ values are estimated in the cases involving transaction costs, which could affect the verification results. In addition, if the verification results are the same in this case then the case where the intercept item equals to zero will also have the same result.

(See Appendix 1.3.2、2.3.2 for the detailed derivations)

The objective function is as following:

$$f(w_1, w_2) = \min_{w_1, w_2} Var\left\langle w_1^T r_1 + w_2^T r_2 \middle| x_0 \right\rangle - \frac{\lambda}{2}\left[E\left\langle w_1^T r_1 + w_2^T r_2 \middle| x_0 \right\rangle - \mu_O\right]$$

The minimum variance of one-stage method is:

$$Var\left\langle w_O^{*\,\mathrm{T}} r \,\middle|\, x_0 \right\rangle = \frac{\mu_O^2}{A^T B^{-1} A}$$

The minimum variance of two-stage method is:

$$Var\left\langle V_\lambda^{*\,T} Z \,\middle|\, x_0 \right\rangle = \frac{\mu_O^2}{F^{*T} G^{*-1} F^*}$$

Before comparing the minimum variance method and the maximum return method, we need to define δ, which denotes the return per unit of risk.

The comparison of the two methods using one-stage method:

The minimum variance method:

$$\delta_{ov} = \frac{\mu_O}{\sqrt{\dfrac{\mu_O^2}{A^T B^{-1} A}}} = \sqrt{A^T B^{-1} A}$$

The maximum return method:

$$\delta_{or} = \frac{\sigma_O \sqrt{A^T B^{-1} A}}{\sigma_O} = \sqrt{A^T B^{-1} A}$$

Above shows that the return per unit of risk remains the same when seeking for a minimum risk for a fixed return using one-stage method.

The comparison of the two methods using two-stage method:

The minimum variance method:

$$\delta_{tv} = \frac{\mu_O}{\sqrt{\frac{\mu_O^2}{F^{*T} G^{*-1} F^*}}} = \sqrt{F^{*T} G^{*-1} F^*}$$

The maximum return method:

$$\delta_{or} = \frac{\sigma_O \sqrt{F^T G^{-1} F}}{\sigma_O} = \sqrt{F^T G^{-1} F}$$

This part can not be proved by the analytical way; we can only verify it with numerical method in the next chapter.

Case #4: Given the condition $\Lambda = 0; \beta = 0$, compare the two dynamic methods to determine which method can give higher expected return after deducting transaction costs. (See Appendix 1.4 、2.4 the detailed derivations)

The objective function is as following:

$$f(w_1, w_2) = \max_{w_1, w_2} E\left\langle w_1^T r_1 + w_2^T r_2 \middle| x_0 \right\rangle - \frac{\lambda}{2}\left[Var\left\langle w_1^T r_1 + w_2^T r_2 \middle| x_0 \right\rangle - \sigma_0^2\right]$$

Expected return of one-stage method is as follows:

$$E\left\langle w_O^{*T} r \middle| x_0 \right\rangle = \sigma_O \sqrt{A_0^T B^{-1} A_0}$$

Expected return of two-stage method is as follows:

$$E\left\langle V_\lambda^{*T} Z \middle| x_0 \right\rangle = \sigma_O \sqrt{F_0^T G_0^{-1} F_0}$$

From section 3.3 and 3.4, we observe that when transaction costs are involved, we can only get estimated λ value whether there is intercept item or not. In contrast, when there is no transaction cost, regardless of the intercept item, λ can always be solved. In addition, in the case $\Lambda = 0; \beta \neq 0$, we can prove that the two methods of Markowitz mean-variance portfolio optimization approach using one-stage method have the same δ value. We will use the numerical methods to verify if the return per unit of risk remains the same when seeking for a minimum risk for a fixed return using two-stage method in the next section.

4 Numerical Illustrations

The data in this paper is collected from Taiwan Economic Journal Data Bank, consisting 253 records of each kind of the Taiwan sector index during 08/18/2009 and 08/18/2010.

4.1 Unit root test and cointegration test

The statistical analysis software Eviews is used to analyze the 19 sector indices, examine if they are stationary, and find the cointegration relation. 5 out of the 19 sector indices from the Taiwan market are chosen for asset allocation, including (1) Building Construction, (2) Financial and Insurance, (3) Steel and Iron, (4) Electronic and Electrical, and (5) Biotech. Data of these 5 sectors are I(1) series. I(1) series, a sequence that can become a stable series after the first difference. Among which (1) Building Construction and (2) Financial and Insurance indices are cointegrated; the other 3 sector indices are not co-integrated with these 2 sectors.

4.2 Parameter estimation

The estimated values for parameters θ_1 、 θ_2 、 Ξ are derived using formula(1):

$$\theta_1 = \begin{pmatrix} 0.6186 \\ 1.5033 \end{pmatrix} ；\ \theta_2 = \begin{pmatrix} 0.0001 \\ 0.0005 \\ 0.0002 \end{pmatrix} ；\ \theta = \begin{pmatrix} 0.6186 \\ 1.5033 \\ 0.0001 \\ 0.0005 \\ 0.0002 \end{pmatrix} ；\ \Xi = \begin{pmatrix} 0.0273 & -0.3392 \\ -0.5222 & -0.6257 \\ 1.3721 & 1.6401 \end{pmatrix}$$

With formula(1) and the estimated θ、Ξ values, use Eviews to estimate the covariance matrix for formula(1)

$$\Sigma = 1.0e-003 * \begin{pmatrix} 0.1044 & 0.0566 & 0.0020 & 0.0001 & -0.0128 \\ 0.0566 & 0.1780 & 0.0144 & 0.0063 & -0.0051 \\ 0.0020 & 0.0144 & 0.0298 & 0.0184 & 0.0190 \\ 0.0001 & 0.0063 & 0.0184 & 0.0291 & 0.0303 \\ -0.0128 & -0.0051 & 0.0190 & 0.0303 & 0.0484 \end{pmatrix}$$

With vector θ and matrix Ξ 、 $\sum$, the estimated values for β 、 Π 、 Π_1 、 $\sum^*$ are calculated from formula (4), (5), and (6).

From formula(4) we get:

$$\beta = \begin{pmatrix} 0.6187 \\ 1.5033 \\ 0.0001 \\ 0.0005 \\ 0.0002 \end{pmatrix}$$

From formula(5) we get:

$$\Pi = \begin{pmatrix} 0 & 0 & 0.0273 & -0.5222 & 1.3721 \\ 0 & 0 & -0.3392 & -0.6257 & 1.6401 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\Pi_1 = \begin{pmatrix} -1 & 0 & 0.0273 & -0.5222 & 1.3721 \\ 0 & -1 & -0.3392 & -0.6257 & 1.6401 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

From formula(6) we get:

$$\Sigma^* = 1.0\mathrm{e}-003 * \begin{pmatrix} 0.1258 & 0.0861 & 0.0193 & 0.0270 & 0.0383 \\ 0.0861 & 0.2131 & 0.0239 & 0.0315 & 0.0489 \\ 0.0193 & 0.0239 & 0.0298 & 0.0184 & 0.0190 \\ 0.0270 & 0.0315 & 0.0184 & 0.0291 & 0.0303 \\ 0.0383 & 0.0489 & 0.0190 & 0.0303 & 0.0484 \end{pmatrix}$$

4.3 Transaction costs

A 0.1425% transaction fee is charged for buying, selling and short selling stocks, and a trading tax of 0.3% is imposed when selling and short selling stocks. Lending fee for short selling is about 0.08%. The above figures do not include any discount. So, in short, the cost of buying stocks is the transaction fee while the cost for selling stock is the trading tax plus the transaction fee; and the cost for short selling is the sum of transaction fee, trading tax, and lending fee. To simplify, a fixed ratio is used as transaction cost for buying, selling and short selling:

(Securities transaction tax + borrowing cost) / 2 + transaction fee = 0.3325%

The transaction cost matrix is as below:

$$\Lambda = \begin{pmatrix} 0.003325 & 0 & 0 & 0 & 0 \\ 0 & 0.003325 & 0 & 0 & 0 \\ 0 & 0 & 0.003325 & 0 & 0 \\ 0 & 0 & 0 & 0.003325 & 0 \\ 0 & 0 & 0 & 0 & 0.003325 \end{pmatrix}$$

4.4 Numerical results

In this section, the real data is applied to the model, and the values of the vectors, matrix, and parameters of the two dynamic programming methods are examined.

4.4.1 Results of one-stage method

The vector and matrix of formula (12) in this paper are:

Expected return matrix A

$$A = \begin{pmatrix} -0.0340 \\ -0.0311 \\ 0.0001 \\ 0.0005 \\ 0.0002 \\ 0.0001 \\ 0.0001 \\ 0.0001 \\ 0.0005 \\ 0.0002 \end{pmatrix}$$

Transaction cost matrix Λ^*

$$\Lambda^* = \begin{pmatrix} 0.00665 & 0 & 0 & 0 & 0 & -0.003325 & 0 & 0 & 0 & 0 \\ 0 & 0.00665 & 0 & 0 & 0 & 0 & -0.003325 & 0 & 0 & 0 \\ 0 & 0 & 0.00665 & 0 & 0 & 0 & 0 & -0.003325 & 0 & 0 \\ 0 & 0 & 0 & 0.00665 & 0 & 0 & 0 & 0 & -0.003325 & 0 \\ 0 & 0 & 0 & 0 & 0.00665 & 0 & 0 & 0 & 0 & -0.003325 \\ -0.003325 & 0 & 0 & 0 & 0 & 0.003325 & 0 & 0 & 0 & 0 \\ 0 & -0.003325 & 0 & 0 & 0 & 0 & 0.003325 & 0 & 0 & 0 \\ 0 & 0 & -0.003325 & 0 & 0 & 0 & 0 & 0.003325 & 0 & 0 \\ 0 & 0 & 0 & -0.003325 & 0 & 0 & 0 & 0 & 0.003325 & 0 \\ 0 & 0 & 0 & 0 & -0.003325 & 0 & 0 & 0 & 0 & 0.003325 \end{pmatrix}$$

Covariance matrix B

$$B = 1.0\mathrm{e}{-003}*$$

$$\begin{pmatrix} 0.1258 & 0.0861 & 0.0193 & 0.0270 & 0.0383 & -0.0868 & -0.0467 & 0 & 0 & 0 \\ 0.0861 & 0.2131 & 0.0239 & 0.0315 & 0.0489 & -0.0349 & -0.1608 & 0 & 0 & 0 \\ 0.0193 & 0.0239 & 0.0298 & 0.0184 & 0.0190 & -0.0020 & -0.0144 & 0 & 0 & 0 \\ 0.0270 & 0.0315 & 0.0184 & 0.0291 & 0.0303 & -0.0001 & -0.0063 & 0 & 0 & 0 \\ 0.0383 & 0.0489 & 0.0190 & 0.0303 & 0.0484 & 0.0128 & 0.0051 & 0 & 0 & 0 \\ -0.0868 & -0.0349 & -0.0020 & -0.0001 & 0.0128 & 0.2301 & 0.1427 & 0.0193 & 0.0270 & 0.0383 \\ -0.0467 & -0.1608 & -0.0144 & -0.0063 & 0.0051 & 0.1427 & 0.3911 & 0.0239 & 0.0315 & 0.0489 \\ 0 & 0 & 0 & 0 & 0 & 0.0193 & 0.0239 & 0.0298 & 0.0184 & 0.0190 \\ 0 & 0 & 0 & 0 & 0 & 0.0270 & 0.0315 & 0.0184 & 0.0291 & 0.0303 \\ 0 & 0 & 0 & 0 & 0 & 0.0383 & 0.0489 & 0.0190 & 0.0303 & 0.0484 \end{pmatrix}$$

4.4.2 Results of two-stage method

The vector and matrix of formula (18) are:

Expected return matrix F

$$F = \begin{pmatrix} -0.0340 \\ -0.0311 \\ 0.0001 \\ 0.0005 \\ 0.0002 \\ 2.1781 \end{pmatrix}$$

Expected transaction cost matrix H

$$H = \begin{pmatrix} 0.0066 & 0 & 0 & 0 & 0 & -0.0055 \\ 0 & 0.0066 & 0 & 0 & 0 & -0.0016 \\ 0 & 0 & 0.0066 & 0 & 0 & -0.0406 \\ 0 & 0 & 0 & 0.0066 & 0 & 0.1358 \\ 0 & 0 & 0 & 0 & 0.0066 & -0.0468 \\ -0.0055 & -0.0016 & -0.0406 & 0.1358 & -0.0468 & 105.5870 \end{pmatrix}$$

Covariance matrix G

$$G = 1.0\text{e} - 003 * \begin{pmatrix} 0.1258 & 0.0861 & 0.0193 & 0.0270 & 0.0383 & 0.3296 \\ 0.0861 & 0.2131 & 0.0239 & 0.0315 & 0.0489 & 0.2676 \\ 0.0193 & 0.0239 & 0.0298 & 0.0184 & 0.0190 & 0.0204 \\ 0.0270 & 0.0315 & 0.0184 & 0.0291 & 0.0303 & 0.0063 \\ 0.0383 & 0.0489 & 0.0190 & 0.0303 & 0.0484 & -0.0469 \\ 0.3296 & 0.2673 & 0.0204 & 0.0063 & -0.0469 & 6869.3291 \end{pmatrix}$$

4.5 Methods Comparison

In order to compare the one-stage method with the two-stage method and determine which method can give higher expected return under the same risk, we apply the estimated parameters and the actual asset price data x_0 to the formula of the expected return and transaction cost derived in this paper.

For the empirical research, we found that the size of Σ will affect the comparison results. We first examine the results by using different Σ values. We will sum the covariance matrix estimated by the 5 sector indices and that the identity matrix multiplied with different coefficients, respectively. And then we observe the changes under the two cases: (1) $\Lambda \neq 0; \beta \neq 0$ (2) $\Lambda = 0; \beta \neq 0$. We only discuss these two cases because the information used in this paper has an intercept item.

According to the first table, the net expected returns of the two-stage method are much better than those of the one-stage method. However, the expected returns of the one-stage method are more sensitive to the change of the covariance matrix. So with the reduction of the diagonal covariance matrix, the net expected returns of the two-stage method will be closer to that of the one-stage method.

Table 1. Comparison of expected net return

Expect net return	$\Lambda \neq 0; \beta \neq 0$	$\Lambda = 0; \beta \neq 0$
Σ +5* identity	(**0.0014**, 0.0408)	(**0.0015**, 0.0408)
Σ +1* identity	(**0.0032**, 0.0409)	(**0.0033**, 0.0409)
Σ +0.5* identity	(**0.0045**, 0.0409)	(**0.0046**, 0.0410)
Σ +0.1* identity	(**0.0097**, 0.0414)	(**0.0103**, 0.0415)
Σ +0.05* identity	(**0.0133**, 0.0419)	(**0.0146**, 0.0421)
Σ +0.01* identity	(**0.0265**, 0.0454)	(**0.0323**, 0.0468)
Σ +0.005* identity	(**0.0345**, 0.0483)	(**0.0453**, 0.0519)
Σ +0.001* identity	(**0.0575**, 0.0583)	(**0.0575**, 0.0583)

Note 1. The results of this table is under $\sigma = 0.05$.

Note 2. The numbers in the table are in the following order: one-stage method, two-stage method.

The second table shows that when the diagonal of covariance matrix is smaller than 10^{-4}, the one-stage method performs better than the two-stage method. The covariance matrix of this paper is about 10^{-4}, so obtained numerical result of the one-stage method is better than that of the two-stage method. In addition, we also observe that when the diagonal of covariance matrix is reduced to 10^{-4} from 10^{-3}, the net expected returns of the above two cases increase significantly.

Table 2. Comparison of expected net return

Expect net return	$\Lambda \neq 0; \beta \neq 0$	$\Lambda = 0; \beta \neq 0$
Σ+0.0001* identity	(**0.1017**, 0.0858)	(**0.1940**, 0.1432)
Σ	(**0.1630**, 0.1249)	(**0.2637**, 0.1916)

Note 1. The results of this table is under $\sigma = 0.05$.

Note 2. The numbers in the table are in the following order: one-stage method, two-stage method.

From the above two tables, we observe that the expected net returns increase with the reduction of the covariance matrix. Relatively, the expected net return of the one-stage method is more sensitive to the change of the covariance matrix than the expected net return of the two-stage method. A slight change to a small covariance matrix can cause significant changes in expected net returns. In addition, when the return rate is lower, there is almost no difference in the expected net returns whether transaction cost is taken into account. When the return rate is higher, the difference between the case of transaction cost and that of no transaction cost is relatively large. However, that won't affect the comparison results of the two methods.

Now we understand the impact of the change of the covariance matrix, let us look at the comparison of two dynamic methods under different standard deviations.

Table 3 shows that the expected net return of the one-stage method is still better than that of the two-stage method under a different standard deviation.

Table 3. Comparison of expected net return

Expect net return	$\sigma_0 = 0.05$	$\sigma_0 = 0.1$	$\sigma_0 = 0.15$
$\Lambda \neq 0; \beta \neq 0$	(**0.1630**, 0.1249)	(**0.2404**, 0.1765)	(**0.2832**, 0.1816)
$\Lambda = 0; \beta \neq 0$	(**0.2637**, 0.1916)	(**0.5274**, 0.3833)	(**0.7911**, 0.5749)

Note. The numbers in the table are in the following order: one-stage method, two-stage method.

In addition, we can see from Table 4 that higher risk gives higher return, but the return per risk unit degrades.

Table 4. Comparison of net return and the return per risk unit under $\Lambda \neq 0; \beta \neq 0$

	$\sigma_0 = 0.05$	$\sigma_0 = 0.1$	$\sigma_0 = 0.15$
Expect net return	(**0.1630**, 0.1249)	(**0.2404**, 0.1765)	(**0.2832**, 0.1816)
Delta	(**3.26**, 2.498)	(**2.404**, 1.916)	(**1.888**, 1.211)

Note: Delta represents the return per risk unit.

Finally, we used the numerical method and verified that the return per unit of risk remains the same when seeking for a minimum risk for a fixed return using two-stage.

5 Conclusions and Recommendations

In the past, many studies conducted in Taiwan and abroad had confirmed the importance of asset allocation. The topic of asset allocation has attracted great attention from investors. Today, with the availability of many investment products, investors can use portfolio to diversify investment risk and apply research results to maximize the return profits.

The shortcomings of the approach of traditional asset allocation are: regardless of the length of the investment, asset allocation decision is made at the beginning of the investment period using the price at that time. The allocation weight will not change over the period of time. The approach of traditional asset allocation does not consider the transaction costs, which is not realistic, and forecast future asset price using past information. In this paper, we extend Markowitz's mean-variance portfolio optimization method to two stages, take into account the transaction costs and then forecast asset price using PCB Cointegration Model. We compared the expected net returns of the two dynamic methods.

Investors will adjust the weight based on latest prices and information using the two-stage method; theoretically it should yield a better expected return. However, according to numerical results from Chapter IV, the expected return of the one-stage method is always better than that of the two-stage method. The reason is that the covariance matrix of the data we chose is very small. If the covariance is very small, the randomness becomes lower, thus the advantage of the two-stage method diminishes.

So it is not unreasonable that the expected return of the one-stage method is better than that of the two-stage method.

In this paper, we extended Markowitz mean-variance portfolio optimization method to a two-period setting. If we adjust the allocation weight in half a year, we can only see the effect of asset allocation about a year long. We recommend future studies on this topic to extend to more periods in order to facilitate long-term investors and frequent traders, and compare the difference with the two-period asset allocation.

Reference

Brinson, G. P., B. D. Singer, and G. L. Beebower (1991), "Determinants of Portfolio Performance: An Update", *Financial Analysts Journal*, Vol. 47, pp. 40-48.

Chopra, V. K. and W. T. Ziemba (1993), "The Effect of Errors in Means, Variances and Covariance on Optimal Portfolio Choice", *Journal of Portfolio Management*, Vol. 19, pp. 6-11.

Elam, E. and B. L. Dixon (1988), "Examining the Validity of a Test of Futures Market Efficiency, *Journal of Futures Markets*, Vol. 8, pp. 365-372

Engle, R. F. and C. W. J. Granger (1987), "Co-integration and Error Correction: Representation, Estimation and Testing," *Econometrica*, Vol. 55, pp. 251-276.

Granger, C. W. J. and P. Newbold (1974), "Spurious Regressions in Economics." *Journal of Econometrics*, Vol. 2, pp. 111-120.

Li, D. and W. L. Ng (2000), "Optimal Dynamic Portfolio Selection: Multiperiod Mean-Variance Formulation", *Mathematical Finance*, Vol. 10, pp. 387-406

Markowitz, H. M. (1952), "Portfolio Selection", *Journal of Finance*, Vol. 7, pp. 77-91.

Melanie, B. R. (2008), "A Dynamic Programming Approach to Two-Stage Mean-Variance Portfolio Selection in Cointegrated Vector Autoregressive Systems", *47th IEEE Conference on Decision and Control Cancun, Mexico, Dec. 9-11.*

Michaud, R. O. (1989), "The Markowitz Optimization Enigma: Is Optimized Optimal?" *Financial Analysts Journal*, Vol. 45, pp. 31-42.

Nelson, C. R. and C. Plosser (1982), "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics*, Vol. 10, pp. 139-167.

Scherer, B. (2002), "Portfolio Resampling: Review and Critique", *Financial Analysts Journal*, Vol. 58, pp. 98-109.

Wolf, M. (2006), "Resampling vs. Shrinkage for Benchmarked Managers", *Working paper,* No. 263, Institute for Empirical Research in Economics University of Zurich, Working Paper Series, ISSN 1424-0459.

Appendix

Appendix 1. One-stage method

Appendix1.1 · $\Lambda \neq 0; \beta \neq 0$

Appendix1.1.1 · Obtained the initial estimate λ_1^* by Taylor expansion

Suppose function $h(\lambda_1)$ is as follows:

$$h(\lambda_1) = Var\left\langle w_{CO}^T r \,\middle|\, x_0 \right\rangle - \sigma_O^2 = A^T \cdot \Psi^{-1} \cdot B \cdot \Psi^{-1} \cdot A - \sigma_0^2$$

$$= A^T \cdot \left(\Lambda^* + \lambda_1 B\right)^{-1} \cdot B \cdot \left(\Lambda^* + \lambda_1 B\right)^{-1} \cdot A - \sigma_0^2$$

where Ψ^{-1} is a function of λ_1, $\Psi^{-1}(\lambda_1) = \left(\Lambda^* + \lambda_1 B\right)^{-1}$.

We have $h(0) = A^T \cdot \Lambda^{*^{-1}} \cdot B \cdot \Lambda^{*^{-1}} \cdot A - \sigma_0^2$

Thm1. Suppose t is a variable in the matrix A and inverse matrix A^{-1} exists, then the first derivative of the inverse matrix A^{-1} is $\frac{\mathrm{dA}^{-1}}{\mathrm{dt}} = -\mathrm{A}^{-1} \cdot \frac{\mathrm{dA}}{\mathrm{dt}} \mathrm{A}^{-1}$.

Proof: $\mathrm{A}^{-1} \cdot \mathrm{A} = \mathrm{I}$; $\frac{\mathrm{dA}^{-1}}{\mathrm{dt}} \cdot \mathrm{A} + \mathrm{A}^{-1} \cdot \frac{\mathrm{dA}}{\mathrm{dt}} = 0$; $\Rightarrow \frac{\mathrm{dA}^{-1}}{\mathrm{dt}} = -\mathrm{A}^{-1} \cdot \frac{\mathrm{dA}}{\mathrm{dt}} \mathrm{A}^{-1}$

By the Thm1, we have:

$$h'(\lambda_1) = -2A^T \cdot \left(\Lambda^* + \lambda_1 B\right)^{-1} \cdot B \cdot \left(\Lambda^* + \lambda_1 B\right)^{-1} \cdot B \cdot \left(\Lambda^* + \lambda_1 B\right)^{-1} \cdot A$$

$$h'(0) = -2A^T \cdot \Lambda^{*^{-1}} \cdot B \cdot \Lambda^{*^{-1}} \cdot B \cdot \Lambda^{*^{-1}} \cdot A$$

Obtain the following estimated formula by Taylor expansion

$h(\lambda_1^*) \approx h(0) + \lambda_1^* h'(0)$ where λ_1^* is the solution of $h(\lambda_1) = Var\left(w_{CO}^T r\right) - \sigma_0^2 = 0$.

$$0 \approx A^T \cdot \Lambda^{*^{-1}} \cdot B \cdot \Lambda^{*^{-1}} \cdot A - \sigma_0^2 - 2\lambda_1^* \cdot A^T \cdot \Lambda^{*^{-1}} \cdot B \cdot \Lambda^{*^{-1}} \cdot B \cdot \Lambda^{*^{-1}} \cdot A$$

Then we get the initial estimate λ_1^* :

$$\lambda_1^* \approx \frac{A^T \cdot \Lambda^{*^{-1}} \cdot B \cdot \Lambda^{*^{-1}} \cdot A - \sigma_0^2}{2A^T \cdot \Lambda^{*^{-1}} \cdot B \cdot \Lambda^{*^{-1}} \cdot B \cdot \Lambda^{*^{-1}} \cdot A}$$

Appendix1.1.2 · The identical equation sort out from the formula $Var\left\langle w_{CO}^T r \,\middle|\, x_0 \right\rangle = \sigma_O^2$

$$Var\left\langle w_{CO}^T r \,\middle|\, x_0 \right\rangle = \sigma_O^2$$

$$A^T \cdot \left(\Lambda^* + \lambda_1 B\right)^{-1} \cdot B \cdot \left(\Lambda^* + \lambda_1 B\right)^{-1} \cdot A = \sigma_0^2$$

$$A^T \cdot \left(\Lambda^* + \lambda_1 B\right)^{-1} \cdot \lambda_1^{-1} \cdot \left(\Lambda^* + \lambda_1 B - \Lambda^*\right) \cdot \left(\Lambda^* + \lambda_1 B\right)^{-1} \cdot A = \sigma_0^2$$

$$A^T \cdot \Psi^{-1} \cdot \left(\Psi - \Lambda^*\right) \cdot \Psi^{-1} \cdot A = \lambda_1 \cdot \sigma_0^2$$

$$A^T \cdot \Psi^{-1} \cdot A - A^T \cdot \Psi^{-1} \cdot \Lambda^* \cdot \Psi^{-1} \cdot A = \lambda_1 \cdot \sigma_0^2$$

Appendix1.2 · $\Lambda \neq 0; \beta = 0$

Objective function for this case is as follows:

$$f_{CO}(w) = E\left\langle w^T r \,\middle|\, x_0 \right\rangle - \frac{1}{2} w^T \Lambda^* w - \frac{\lambda_3}{2}\left[Var\left\langle w^T r \,\middle|\, x_0 \right\rangle - \sigma_0^2\right]$$
$$= w^T A_0 - \frac{1}{2} w^T \Lambda^* w - \frac{\lambda_3}{2}\left[w^T B w - \sigma_0^2\right]$$

The vectors and matrices represent respectively:

$$w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \ ; \qquad r = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \ ; \qquad \Lambda^* = \begin{bmatrix} 2\Lambda & -\Lambda \\ -\Lambda & \Lambda \end{bmatrix}$$

$$A_0 = E\langle r \,|\, x_0 \rangle = \begin{bmatrix} E\langle r_1 \,|\, x_0 \rangle \\ E\langle r_2 \,|\, x_0 \rangle \end{bmatrix} \ ;$$

$$B = Var\langle r \,|\, x_0 \rangle = \begin{bmatrix} Var\langle r_1 \,|\, x_0 \rangle & Cov\langle r_1, r_2 \,|\, x_0 \rangle \\ Cov\langle r_2, r_1 \,|\, x_0 \rangle & Var\langle r_2 \,|\, x_0 \rangle \end{bmatrix}$$

To optimize the objective function, we do the first order differential equal to zero on it.

$$A_0 - \Lambda^* w - \lambda_3 B w = 0$$

$$w_{CO} = \Psi^{-1} A_0 \qquad \text{where} \quad \Psi = \left(\Lambda^* + \lambda_3 B\right)$$

According to Eq.8 and Eq.9 in section 3.1, the derived results are as follows:

$$E\langle r_1 \,|\, x_0 \rangle = E\langle \Pi_1 x_0 + u_1 \,|\, x_0 \rangle = \Pi_1 x_0$$

$$E\langle r_2 \,|\, x_0 \rangle = E\langle \Pi_1 u_1 + u_2 \,|\, x_0 \rangle = 0$$

$$Var\langle r_1 \,|\, x_0 \rangle = Var\langle \Pi_1 x_0 + u_1 \,|\, x_0 \rangle = \Sigma^*$$

$$Var\langle r_2 \,|\, x_0 \rangle = Var\langle \Pi_1 u_1 + u_2 \,|\, x_0 \rangle = D \qquad \text{where} \quad D = \Pi_1 \Sigma^* \Pi_1^T + \Sigma^*$$

$$Cov\langle r_1, r_2 \,|\, x_0 \rangle = Cov\langle \Pi_1 x_0 + u_1, \Pi_1 u_1 + u_2 \,|\, x_0 \rangle$$

$$= Cov\langle u_1, \Pi_1 u_1 \,|\, x_0 \rangle = \Sigma^* \Pi_1^T$$

According to the derived results, we understand that, in this case, only the matrix of the expected return will be changed to matrix A_0. Then following the same approach as before, we get the expected return after deducting transaction cost of one-stage method as follows:

$$E\left\langle {w_{CO}^*}^T r \,\middle|\, x_0 \right\rangle - \frac{1}{2} {w_{CO}^*}^T \Lambda^* w_{CO}^* = \frac{1}{2}\left(A_0^T \cdot \Psi^{*-1} \cdot A_0 + \lambda_3^S \cdot \sigma_0^2\right) \qquad \text{where} \quad \Psi^* = \left(\Lambda^* + \lambda_3^S B\right)$$

Appendix1.3、 $\Lambda = 0; \beta \neq 0$

Appendix1.3.1、One-stage method of the maximum return method

Objective function for this case is as follows:

$$f_O(w) = E\left\langle w^T r \middle| x_0 \right\rangle - \frac{\lambda_5}{2}\left[Var\left\langle w^T r \middle| x_0 \right\rangle - \sigma_0^2\right]$$
$$= w^T A - \frac{\lambda_5}{2}\left[w^T B w - \sigma_0^2\right]$$

The vectors and matrices represent respectively:

$$w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \;;\quad r = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \;;\quad A = E\langle r | x_0 \rangle = \begin{bmatrix} E\langle r_1 | x_0 \rangle \\ E\langle r_2 | x_0 \rangle \end{bmatrix} \;;$$

$$B = Var\langle r | x_0 \rangle = \begin{bmatrix} Var\langle r_1 | x_0 \rangle & Cov\langle r_1, r_2 | x_0 \rangle \\ Cov\langle r_2, r_1 | x_0 \rangle & Var\langle r_2 | x_0 \rangle \end{bmatrix}$$

To optimize objective function, we do the first order differential equal to zero on it.

$$A - \lambda_5 B w = 0$$

$$w_O = \frac{1}{\lambda_5} B^{-1} A$$

The vector A and the matrix B here are the same as described earlier in this paper.

Use the formula, $Var\left\langle w_O^T r \middle| x_0 \right\rangle = \sigma_O^2$, to solve λ_5.

$$\sigma_O^2 = Var\left\langle w_O^T r \middle| x_0 \right\rangle = w_O^T Var\langle r | x_0 \rangle w_O = \frac{1}{\lambda_5^2} A^T B^{-1} B B^{-1} A = \frac{1}{\lambda_5^2} A^T B^{-1} A$$

$$\lambda_5 = \frac{\sqrt{A^T B^{-1} A}}{\sigma_O}$$

The optimum weight is: $w_O^* = \dfrac{\sigma_O}{\sqrt{A^T B^{-1} A}} B^{-1} A$

The expected return is: $E\left\langle w_O^{*\,T} r \middle| x_0 \right\rangle = \dfrac{\sigma_O}{\sqrt{A^T B^{-1} A}} \cdot A^T B^{-1} A = \sigma_O \sqrt{A^T B^{-1} A}$

Next we will switch the method from maximizing the return to minimizing the variance of mean-variance portfolio optimization approach and discuss the performance of the one-stage method.

Appendix1.3.2 · One-stage method of the minimum variance method

The objective function can be easily written down as follows:

$$f_O(w) = Var\left\langle w^T r \middle| x_0 \right\rangle - \frac{\lambda_{v1}}{2}\left[E\left\langle w^T r \middle| x_0 \right\rangle - \mu_O\right]$$
$$= w^T Bw - \frac{\lambda_{v1}}{2}\left[w^T A - \mu_O\right]$$

Do the first order differential equal to zero on it.

$$2Bw - \frac{\lambda_{v1}}{2} A = 0$$
$$w_O = \frac{\lambda_{v1}}{4} B^{-1} A$$

Here the vector A and the matrix B are the same as described earlier in the paper.

Use the formula, $E\left(w_O^T r\right) = \mu_O$, to solve λ_{1v}.

$$\mu_O = E\left(w_O^{*\,\mathrm{T}} r\right) = \frac{\lambda_{1v}}{4} A^T B^{-1} A$$

$$\lambda_{1v} = \frac{4\mu_O}{A^T B^{-1} A}$$

The optimum weight is: $w_O^* = \dfrac{\mu_O}{A^T B^{-1} A} \cdot B^{-1} A$

The minimum variance is: $Var\left\langle w_O^{*\,\mathrm{T}} r \middle| x_0 \right\rangle = w_O^{*\,\mathrm{T}} Var\left\langle r \middle| x_0 \right\rangle w_O^* = \dfrac{\mu_O^2}{A^T B^{-1} A}$

Appendix1.4 · $\Lambda = 0; \beta = 0$

Objective function for this case is as follows:

$$f_O(w) = E\left\langle w^T r \middle| x_0 \right\rangle - \frac{\lambda_7}{2}\left[Var\left\langle w^T r \middle| x_0 \right\rangle - \sigma_0^2\right]$$

$$= w^T A_0 - \frac{\lambda_7}{2}\left[w^T B w - \sigma_0^2\right]$$

The vectors and matrices represent respectively:

$$w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \ ; \qquad r = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \ ; \qquad A_0 = E\langle r \,|\, x_0 \rangle = \begin{bmatrix} E\langle r_1 \,|\, x_0 \rangle \\ E\langle r_2 \,|\, x_0 \rangle \end{bmatrix} \ ;$$

$$B = Var\langle r \,|\, x_0 \rangle = \begin{bmatrix} Var\langle r_1 \,|\, x_0 \rangle & Cov\langle r_1, r_2 \,|\, x_0 \rangle \\ Cov\langle r_2, r_1 \,|\, x_0 \rangle & Var\langle r_2 \,|\, x_0 \rangle \end{bmatrix}$$

Do the first order differential equal to zero on it.

$$w_O = \frac{1}{\lambda_7} B^{-1} A_0$$

Here the vector A_0 and the matrix B are the same as described earlier in the paper.

Then solve λ_7 and get the expected return as follows:

$$E\left\langle {w_O^*}^T r \,\middle|\, x_0 \right\rangle = \sigma_O \sqrt{A_0^T B^{-1} A_0}$$

Appendix 2. Two-stage method

Appendix2.1 · $\Lambda \neq 0; \beta \neq 0$

Appendix2.1.1 · Deriving the vector of expected return and the matrices of variance and expected transaction cost

Before discussing the expected return of the second phase, let take a look at two theorems first.

Thm2. x is a random vector of dimension $n \times 1$, $x \sim N(\mu, \Sigma)$ and A is a symmetric matrix, then $E(x^T A x) = trace(A\Sigma)$.

Thm3. $\Phi^{\mathrm{T}}\Pi_1$ is a symmetric matrix.

Proof: We know $\Pi_1 = \Pi - I = \begin{bmatrix} -I & \Xi^T \\ 0 & 0 \end{bmatrix}$; $\Pi_1^T = \begin{bmatrix} -I & 0 \\ \Xi & 0 \end{bmatrix}$

Suppose $\left(\Sigma^*\right)^{-1} = \begin{bmatrix} a_1 & a_2 \\ a_2^T & a_3 \end{bmatrix}$ is a symmetric matrix. (where $a_1^T = a_1$)

Then $\Phi^{\mathrm{T}}\Pi_1 = \Pi_1^{\mathrm{T}}\left(\Sigma^*\right)^{-1}\Pi_1 = \begin{bmatrix} a_1 & -a_1\Xi^T \\ -\Xi a_1 & \Xi a_1\Xi^T \end{bmatrix}$ is a symmetric matrix.

By Thm2. and Thm3, the expected return of the second stage is:

$$E\left\langle {w_2^*}^T r_2 \,\middle|\, x_0 \right\rangle$$

$$= E\left\langle \left(\gamma^T + x_1^T\Phi^{\mathrm{T}}\right) r_2 \,\middle|\, x_0 \right\rangle$$

$$= E\left\langle \gamma^T r_2 + x_1^T\Phi^{\mathrm{T}} r_2 \,\middle|\, x_0 \right\rangle$$

$$= E\left\langle \gamma^T\left(\Pi\beta + \Pi_1 u_1 + u_2\right) + \left(\beta^T + x_0^T\Pi^T + u_1^T\right)\Phi^{\mathrm{T}}\left(\Pi\beta + \Pi_1 u_1 + u_2\right) \,\middle|\, x_0 \right\rangle$$

$$= \gamma^T\Pi\beta + \beta^T\Phi^{\mathrm{T}}\Pi\beta + tr(\varphi)$$

$= \ell$ where $\ell = L^T\Pi\beta + tr(\varphi)$; $L = (\gamma + \Phi\beta)$; $\varphi = \Phi^{\mathrm{T}}\Pi_1\Sigma^*$

Before discussing the variance of the second phase, we first take a look at the two theorems.

Thm4. x is a random vector of dimension $n\times 1$, $x \sim N(\mu,\Sigma)$, and A is a symmetric matrix, then $Var\left(x^T Ax\right) = 2tr(A\Sigma A\Sigma)$.

Thm5. $\mathrm{V}ar\left(u_1^T \Phi^{\mathrm{T}} u_2\right) = tr\left(\Phi^{\mathrm{T}} \Pi_1 \Sigma^*\right)$

Proof: $\mathrm{V}ar\left(u_1^T \Phi^{\mathrm{T}} u_2\right)$

$$= E\left[\left(u_1^T \Phi^{\mathrm{T}} u_2\right)^2\right] - \left[E\left(u_1^T \Phi^{\mathrm{T}} u_2\right)\right]^2$$

$$= E\left[\left(u_1^T \Phi^{\mathrm{T}} u_2\right)^2\right] = E\left[u_1^T \Phi^{\mathrm{T}} u_2 \cdot u_1^T \Phi^{\mathrm{T}} u_2\right] = E\left[u_1^T \Phi^{\mathrm{T}} u_2 \cdot u_2^T \Phi u_1\right]$$

$$= E\left[u_1^T E\left\langle \Phi^{\mathrm{T}} u_2 \cdot u_2^T \Phi \middle| u_1 \right\rangle u_1\right] = E\left(u_1^T \Phi^{\mathrm{T}} \Sigma^* \Phi u_1\right) = E\left[tr\left(u_1^T \Phi^{\mathrm{T}} \Sigma^* \Phi u_1\right)\right]$$

$$= E\left[tr\left(\Phi^{\mathrm{T}} \Sigma^* \Phi u_1 u_1^T\right)\right] = tr\left(\Phi^{\mathrm{T}} \Pi_1 \Sigma^*\right)$$

By Thm4. and Thm5, the variance of the second stage is:

$$Var\left\langle w_2^{*T} r_2 \middle| x_0 \right\rangle$$

$$= \mathrm{V}ar\left\langle \left(\gamma^T + x_1^T \Phi^{\mathrm{T}}\right)\left(\Pi\beta + \Pi_1 u_1 + u_2\right) \middle| x_0 \right\rangle$$

$$= \mathrm{V}ar\left\langle \gamma^T \left(\Pi\beta + \Pi_1 u_1 + u_2\right) \middle| x_0 \right\rangle \tag{A-1}$$

$$+ \mathrm{V}ar\left\langle x_1^T \Phi^{\mathrm{T}} \left(\Pi\beta + \Pi_1 u_1 + u_2\right) \middle| x_0 \right\rangle \tag{A-2}$$

$$+ 2\gamma^T \mathrm{Cov}\left\langle \left(\Pi_1 u_1 + u_2\right), x_1^T \Phi^{\mathrm{T}} \left(\Pi\beta + \Pi_1 u_1 + u_2\right) \middle| x_0 \right\rangle \tag{A-3}$$

$$= k$$

where $D = \Pi_1 \Sigma^* \Pi_1^T + \Sigma^*$; $\varphi = \Phi^{\mathrm{T}} \Pi_1 \Sigma^*$; $L = \left(\gamma + \Phi\beta\right)$;

$$k = \gamma^T D\gamma + \left(\gamma^T + L^{\mathrm{T}}\right) D\Phi\beta + 2tr\left(\varphi^2\right) + tr\left(\varphi\right) + \left(\beta^T \Pi^T \Phi + 2L^T \Pi_1\right)\Sigma^* \Phi^{\mathrm{T}} \Pi\beta$$

Detail in each part, respectively:

(A-1)： $\mathrm{V}ar\left\langle \gamma^T \left(\Pi\beta + \Pi_1 u_1 + u_2\right) \middle| x_0 \right\rangle$

$$= \gamma^T \left(\Pi_1 \Sigma^* \Pi_1^T + \Sigma^*\right)\gamma$$

$$= \gamma^T D\gamma$$

(A-2)： $\mathrm{V}ar\left\langle x_1^T \Phi^{\mathrm{T}} (\Pi\beta + \Pi_1 u_1 + u_2) \middle| x_0 \right\rangle$

$$= \mathrm{V}ar\left\langle \left(\beta^T + x_0^T \Pi^T + u_1^T\right) \Phi^{\mathrm{T}} (\Pi\beta + \Pi_1 u_1 + u_2) \middle| x_0 \right\rangle$$

$$= \beta^T \Phi^{\mathrm{T}} \left(\Pi_1 \Sigma^* \Pi_1^T + \Sigma^*\right) \Phi\beta + \beta^T \Pi^T \Phi \Sigma^* \Phi^{\mathrm{T}} \Pi\beta + 2tr\left(\Phi^{\mathrm{T}} \Pi_1 \Sigma^* \Phi^{\mathrm{T}} \Pi_1 \Sigma^*\right)$$

$$+ tr\left(\Phi^{\mathrm{T}} \Pi_1 \Sigma^*\right) + 2\beta^T \Phi^{\mathrm{T}} \Pi_1 \Sigma^* \Phi^{\mathrm{T}} \Pi\beta$$

$$= \beta^T \Phi^{\mathrm{T}} D \Phi\beta + 2tr\left(\varphi^2\right) + tr(\varphi) + \left(\beta^T \Pi^T \Phi + 2\beta^T \Phi^{\mathrm{T}} \Pi_1\right) \Sigma^* \Phi^{\mathrm{T}} \Pi\beta$$

(A-3)： $2\gamma^T \mathrm{Cov}\left\langle (\Pi_1 u_1 + u_2), x_1^T \Phi^{\mathrm{T}} (\Pi\beta + \Pi_1 u_1 + u_2) \middle| x_0 \right\rangle$

$$= 2\gamma^T \mathrm{Cov}\left\langle \Pi_1 u_1, \left(\beta^T + x_0^T \Pi^T + u_1^T\right) \Phi^{\mathrm{T}} (\Pi\beta + \Pi_1 u_1 + u_2) \middle| x_0 \right\rangle$$

$$+ 2\gamma^T \mathrm{Cov}\left\langle u_2, \left(\beta^T + x_0^T \Pi^T + u_1^T\right) \Phi^{\mathrm{T}} (\Pi\beta + \Pi_1 u_1 + u_2) \middle| x_0 \right\rangle$$

$$= 2\gamma^T \Pi_1 \Sigma^* \Pi_1^T \Phi\beta + 2\gamma^T \Pi_1 \Sigma^* \Phi^{\mathrm{T}} \Pi\beta + 2\gamma^T \Sigma^* \Phi\beta$$

$$= 2\gamma^T \left(\Pi_1 \Sigma^* \Pi_1^T + \Sigma^*\right) \Phi\beta + 2\gamma^T \Pi_1 \Sigma^* \Phi^{\mathrm{T}} \Pi\beta$$

$$= 2\gamma^T D \Phi\beta + 2\gamma^T \Pi_1 \Sigma^* \Phi^{\mathrm{T}} \Pi\beta$$

The covariance of two stages:

$$Cov\left\langle r_1, w_2^{*T} r_2 \middle| x_0 \right\rangle$$

$$= Cov\left\langle \beta + \Pi_1 x_0 + u_1, \left(\gamma^T + x_1^T \Phi^{\mathrm{T}}\right) (\Pi\beta + \Pi_1 u_1 + u_2) \middle| x_0 \right\rangle$$

$$= Cov\left\langle u_1, \gamma^T (\Pi\beta + \Pi_1 u_1 + u_2) + \left(\beta^T + x_0^T \Pi^T + u_1^T\right) \Phi^{\mathrm{T}} (\Pi\beta + \Pi_1 u_1 + u_2) \middle| x_0 \right\rangle$$

$$= Cov\left\langle u_1, \gamma^T \Pi_1 u_1 + \beta^T \Phi^{\mathrm{T}} \Pi_1 u_1 + u_1^T \Phi^{\mathrm{T}} \Pi\beta \middle| x_0 \right\rangle$$

$$= N \qquad \text{where} \quad N = \Sigma^* \left(\Pi_1^T \gamma + \Pi_1^T \Phi\beta + \Phi^{\mathrm{T}} \Pi\beta\right)$$

Let us take a look at another theorem before we discuss transaction cost of two stages.

Thm6. $\Phi^{\mathrm{T}}\Lambda\Phi$ is a symmetric matrix.

Proof: Suppose $\left(\Sigma^*\right)^{-1} = \begin{bmatrix} a_1 & a_2 \\ a_2^T & a_3 \end{bmatrix}$ and $\Lambda = \begin{bmatrix} \Lambda_1 & \Lambda_2 \\ \Lambda_2^T & \Lambda_3 \end{bmatrix}$ are symmetric matrices.

Then $\left(\Sigma^*\right)^{-1}\Lambda\left(\Sigma^*\right)^{-1}$ is a symmetric matrix.

By Thm3. $\Phi^{\mathrm{T}}\Lambda\Phi = \Pi_1^{\mathrm{T}}\left(\Sigma^{-1}\Lambda\Sigma^{-1}\right)\Pi_1$ is a symmetric matrix.

Transaction cost of two stages:

$$E\left\langle 2\Lambda \middle| x_0 \right\rangle = 2\Lambda$$

$$E\left\langle \Lambda w_2^* \middle| x_0 \right\rangle$$

$$= \Lambda E\left\langle \gamma + \Phi x_1 \middle| x_0 \right\rangle$$

$$= \Lambda E\left\langle \gamma + \Phi\left(\beta + \Pi x_0 + u_1\right) \middle| x_0 \right\rangle$$

$$= \Lambda L \qquad \text{where} \quad L = \left(\gamma + \Phi\beta\right)$$

$$E\left\langle w_2^{*T}\Lambda w_2^* \middle| x_0 \right\rangle$$

$$= E\left\langle \left(\gamma^T + x_1^T\Phi^{\mathrm{T}}\right)\Lambda\left(\gamma + \Phi x_1\right) \middle| x_0 \right\rangle$$

$$= E\left\langle \left[\gamma^T + \left(\beta^T + x_0^T\Pi^T + u_1^T\right)\Phi^{\mathrm{T}}\right]\Lambda\left[\gamma + \Phi\left(\beta + \Pi x_0 + u_1\right)\right] \middle| x_0 \right\rangle$$

$$= \gamma^T\Lambda\gamma + \gamma^T\Lambda\Phi\beta + \beta^T\Phi^{\mathrm{T}}\Lambda\gamma + \beta^T\Phi^{\mathrm{T}}\Lambda\Phi\beta + E\left\langle u_1^T\Phi^{\mathrm{T}}\Lambda\Phi u_1 \middle| x_0 \right\rangle$$

$$= L^T\Lambda L + tr\left(\varphi_1\right) \quad \text{where} \quad L = \left(\gamma + \Phi\beta\right) \text{；} \quad \varphi_1 = \Phi^{\mathrm{T}}\Lambda\Phi\Sigma^*$$

Appendix 2.1.2 · Obtained the initial estimate λ_2^* by Taylor expansion

Suppose function $h\left(\lambda_2\right)$ is as follows:

$$h\left(\lambda_2\right) = Var\left\langle V^{*T}Z \middle| x_0 \right\rangle - \sigma_0^2$$

$$= F^T \cdot \Omega^{-1} \cdot G \cdot \Omega^{-1} \cdot F - \sigma_0^2$$

$$= F^T \cdot \left(H + \lambda_2 G\right)^{-1} \cdot G \cdot \left(H + \lambda_2 G\right)^{-1} \cdot F - \sigma_0^2$$

where Ω^{-1} is a function of λ_2, $\Omega^{-1}(\lambda_2) = \left(H + \lambda_2 G\right)^{-1}$.

We have $h(0) = F^T \cdot H^{-1} \cdot G \cdot H^{-1} \cdot F - \sigma_0^2$

By Thm1, we know $h'(\lambda_2) = -2F^T \cdot \left(H + \lambda_2 G\right)^{-1} \cdot G \cdot \left(H + \lambda_2 G\right)^{-1} \cdot G \cdot \left(H + \lambda_2 G\right)^{-1} \cdot F$

$$h'(0) = -2F^T \cdot H^{-1} \cdot G \cdot H^{-1} \cdot G \cdot H^{-1} \cdot F$$

We can obtain the following estimated formula by Taylor expansion:

$h(\lambda_2^*) \approx h(0) + \lambda_2^* h'(0)$ where λ_2^* is the solution of $h(\lambda_2) = Var\left(V^{*T} r\right) - \sigma_0^2 = 0$.

$$0 \approx F^T \cdot H^{-1} \cdot G \cdot H^{-1} \cdot F - \sigma_0^2 - 2\lambda_2^* \cdot F^T \cdot H^{-1} \cdot G \cdot H^{-1} \cdot G \cdot H^{-1} \cdot F$$

Then we get the initial estimate λ_2^*:

$$\lambda_2^* \approx \frac{F^T \cdot H^{-1} \cdot G \cdot H^{-1} \cdot F - \sigma_0^2}{2F^T \cdot H^{-1} \cdot G \cdot H^{-1} \cdot G \cdot H^{-1} \cdot F}$$

Appendix2.1.3 · The identical equation sort out from the formula $Var\left\langle V^{*T} Z \middle| x_0 \right\rangle = \sigma_0^2$

$$Var\left\langle V^{*T} Z \middle| x_0 \right\rangle = \sigma_0^2$$

$$F^T \cdot \left(H + \lambda_2 G\right)^{-1} \cdot G \cdot \left(H + \lambda_2 G\right)^{-1} \cdot F = \sigma_0^2$$

$$F^T \cdot \left(H + \lambda_2 G\right)^{-1} \cdot \lambda_2^{-1} \cdot \left(H + \lambda_2 G - H\right) \cdot \left(H + \lambda_2 G\right)^{-1} \cdot F = \sigma_0^2$$

$$F^T \cdot \Omega^{-1} \cdot \left(\Omega - H\right) \cdot \Omega^{-1} \cdot F = \lambda_2 \cdot \sigma_0^2$$

$$F^T \cdot \Omega^{-1} \cdot F - F^T \cdot \Omega^{-1} \cdot H \cdot \Omega^{-1} \cdot F = \lambda_2 \cdot \sigma_0^2$$

So far we observe that under the assumption in this paper, the optimum weight of the second phase is not relevant to the transaction cost. If Λ changes, it will not affect its results; if β changes, it will. When $\beta \neq 0$, as described in the paper, $w_2^* = \gamma + \Phi x_1$ where $\gamma = (\Sigma^*)^{-1}\beta$, $\Phi = (\Sigma^*)^{-1}\Pi_1$, and if you give a coefficient, a_1, then it becomes $a_1 w_2^* = a_1(\gamma + \Phi x_1)$;when $\beta = 0$,then $w_2^* = \Phi x_1$ where $\Phi = (\Sigma^*)^{-1}\Pi_1$, and if you give a coefficient, a_1, then it becomes $a_1 w_2^* = a_1 \Phi x_1$. In the following verification, we will use this result directly.

Appendix2.2、 $\Lambda \neq 0; \beta = 0$

When x_0, x_1 are knew, the weight of the second stage is $w_2^* = \Phi x_1$ where $(\Sigma^*)^{-1}\Pi_1 = \Phi$. While given a coefficient, a_1, it becomes $a_1 w_2^* = a_1 \Phi x_1$. Then apply the optimal weight of the second phase to the optimal objective function of two periods, as shown below:

$$f_{CT}(V) = E\left\langle V^T Z \middle| x_0 \right\rangle - \frac{1}{2} E\left\langle V^T C V \middle| x_0 \right\rangle - \frac{\lambda_4}{2}\left[Var\left\langle V^T Z \middle| x_0 \right\rangle - \sigma_0^2\right]$$
$$= V^T F_0 - \frac{1}{2} V^T H_0 V - \frac{\lambda_4}{2}\left[V^T G_0 V - \sigma_0^2\right]$$

The vectors and matrices represent respectively:

$$V = \begin{bmatrix} w_1 \\ a_1 \end{bmatrix} \; ; \quad Z = \begin{bmatrix} r_1 \\ w_2^{*T} r_2 \end{bmatrix} \; ; \quad C = \begin{bmatrix} 2\Lambda & -\Lambda w_2^* \\ -w_2^{*T}\Lambda & w_2^{*T}\Lambda w_2^* \end{bmatrix} \; ;$$

$$F_0 = E\langle Z | x_0 \rangle = \begin{bmatrix} E\langle r_1 | x_0 \rangle \\ E\left\langle w_2^{*T} r_2 \middle| x_0 \right\rangle \end{bmatrix} \; ;$$

$$H_0 = E\langle C | x_0 \rangle = \begin{bmatrix} E\langle 2\Lambda | x_0 \rangle & -E\left\langle \Lambda w_2^* \middle| x_0 \right\rangle \\ -E\left\langle w_2^{*T}\Lambda \middle| x_0 \right\rangle & E\left\langle w_2^{*T}\Lambda w_2^* \middle| x_0 \right\rangle \end{bmatrix}$$

$$G_0 = Var\left\langle Z \,\middle|\, x_0 \right\rangle = \begin{bmatrix} Var\left\langle r_1 \,\middle|\, x_0 \right\rangle & Cov\left\langle r_1, w_2^{*T} r_2 \,\middle|\, x_0 \right\rangle \\ Cov\left\langle w_2^{*T} r_2, r_1 \,\middle|\, x_0 \right\rangle & Var\left\langle w_2^{*T} r_2 \,\middle|\, x_0 \right\rangle \end{bmatrix}$$

After taking the first order differential, we get

$$F_0 - H_0 V - \lambda_4 G_0 V = 0$$

$$V^* = \Omega_0^{\ -1} F_0 \quad \text{where} \ \ \Omega_0 = \left(H_0 + \lambda_4 G_0\right)$$

According to Eq.8 and Eq.9 in section 3.1, the derived results are as follows:

$$E(r_1) = E(\Pi_1 x_0 + u_1) = \Pi_1 x_0$$

$$E\left\langle w_2^{*T} r_2 \,\middle|\, x_0 \right\rangle = E\left\langle x_1^T \Phi^{\mathrm{T}} r_2 \,\middle|\, x_0 \right\rangle = tr(\varphi) \qquad \text{where}\, \varphi = \Phi^{\mathrm{T}} \Pi_1 \Sigma^*$$

$$Var\left\langle r_1 \,\middle|\, x_0 \right\rangle = Var\left\langle \Pi_1 x_0 + u_1 \,\middle|\, x_0 \right\rangle = \Sigma^*$$

$$Var\left\langle w_2^{*T} r_2 \,\middle|\, x_0 \right\rangle = Var\left\langle x_1^T \Phi^{\mathrm{T}} r_2 \,\middle|\, x_0 \right\rangle = 2tr\left(\varphi^2\right) + tr(\varphi) \quad \text{where}\, \varphi = \Phi^{\mathrm{T}} \Pi_1 \Sigma^*$$

$$Cov\left\langle r_1, w_2^{*T} r_2 \,\middle|\, x_0 \right\rangle = Cov\left\langle \Pi_1 x_0 + u_1, x_1^T \Phi^{\mathrm{T}} \left(\Pi_1 u_1 + u_2\right) \middle|\, x_0 \right\rangle = 0$$

$$E\left\langle 2\Lambda \,\middle|\, x_0 \right\rangle = 2\Lambda$$

$$E\left\langle \Lambda w_2^* \,\middle|\, x_0 \right\rangle = 0$$

$$E\left\langle w_2^{*T} \Lambda w_2^* \,\middle|\, x_0 \right\rangle = tr(\varphi_1) \quad \text{where}\, \varphi_1 = \Phi^{\mathrm{T}} \Lambda \Phi \Sigma^*$$

From the derived results, we learn that, in this case, the matrices of the expected return, variance, and transaction cost will be changed into matrix F_0 、 H_0 、 G_0 respectively. Then following the same approach as before, we can get the expected return after deducting transaction cost of two-stage method as follows:

$$E\left\langle V_\lambda^{*T} Z \,\middle|\, x_0 \right\rangle - \frac{1}{2} E\left\langle V_\lambda^{*T} C V_\lambda^* \,\middle|\, x_0 \right\rangle = \frac{1}{2}\left(F_0^T \cdot \Omega_0^{*-1} \cdot F_0 + \lambda_4^S \cdot \sigma_0^2\right) \ \text{where} \ \ \Omega_0^* = \left(H + \lambda_4^S G\right)$$

Appendix2.3 · $\Lambda = 0; \beta \neq 0$

Appendix2.3.1 · Two-stage method of the maximum return method

When x_0, x_1 are knew, the weight of the second stage is $w_2^* = \gamma + \Phi x_1$ where $(\Sigma^*)^{-1}\beta = \gamma$; $(\Sigma^*)^{-1}\Pi_1 = \Phi$. While given a coefficient, a_1 , it becomes $a_1 w_2^* = a_1(\gamma + \Phi x_1)$. Then apply the optimal weight of the second phase to the optimal objective function of the two-stage method as shown below:

$$
\begin{aligned}
f_T(V) &= E\left\langle V^T Z \,\middle|\, x_0 \right\rangle - \frac{\lambda_6}{2}\left[Var\left\langle V^T Z \,\middle|\, x_0 \right\rangle - \sigma_0^2\right] \\
&= V^T F - \frac{\lambda_6}{2}\left[V^T G V - \sigma_0^2\right]
\end{aligned}
$$

The vectors and matrices represent respectively:

$$
V = \begin{bmatrix} w_1 \\ a_1 \end{bmatrix} ; \quad Z = \begin{bmatrix} r_1 \\ w_2^{*T} r_2 \end{bmatrix} ; \quad F = E\langle Z \mid x_0 \rangle = \begin{bmatrix} E\langle r_1 \mid x_0 \rangle \\ E\left\langle w_2^{*T} r_2 \,\middle|\, x_0 \right\rangle \end{bmatrix} ;
$$

$$
G = Var\langle Z \mid x_0 \rangle = \begin{bmatrix} Var\langle r_1 \mid x_0 \rangle & Cov\left\langle r_1, w_2^{*T} r_2 \,\middle|\, x_0 \right\rangle \\ Cov\left\langle w_2^{*T} r_2, r_1 \,\middle|\, x_0 \right\rangle & Var\left\langle w_2^{*T} r_2 \,\middle|\, x_0 \right\rangle \end{bmatrix}
$$

After taking the first order differential, we get

$$
F - \lambda_6 G V = 0
$$

$$
V^* = \frac{1}{\lambda_6} G^{-1} F
$$

Here the vector F and the matrix G are the same as described earlier in the paper.

Use the formula, $Var\left\langle V^{*T} Z \,\middle|\, x_0 \right\rangle = \sigma_O^2$, to solve λ_6.

$$
\sigma_O^2 = Var\left\langle V^{*T} Z \,\middle|\, x_0 \right\rangle = V^{*T} Var\langle Z \mid x_0 \rangle V^* = \frac{1}{\lambda_6^2} F^T G^{-1} G G^{-1} F = \frac{1}{\lambda_6^2} F^T G^{-1} F
$$

$$\lambda_6 = \frac{\sqrt{F^T G^{-1} F}}{\sigma_O}$$

The vector V_λ^* that represents the combination of the optimal weight of the first stage and an optimal factor of the second stage:

$$V_\lambda^* = \frac{\sigma_O}{\sqrt{F^T G^{-1} F}} G^{-1} F$$

The expected return of two-stage method as follows:

$$E\left\langle V_\lambda^{*T} Z \,\middle|\, x_0 \right\rangle = \frac{\sigma_O}{\sqrt{F^T G^{-1} F}} \cdot F^T G^{-1} F = \sigma_O \sqrt{F^T G^{-1} F}$$

Next we will switch the method from maximizing the return to minimizing the variance of mean-variance portfolio optimization approach and discuss the performance of the two-stage method.

Appendix2.3.2 · Two-stage method of the minimum variance method

Suppose x_0 and x_1 are known. Then the objective function of the second phase is as follows:

$$\begin{aligned} f_2(w_2) &= Var\left\langle w_2^T r_2 \,\middle|\, x_0, x_1 \right\rangle - \frac{1}{2} E\left\langle w_2^T r_2 \,\middle|\, x_0, x_1 \right\rangle \\ &= w_2^T Var\left\langle r_2 \,|\, x_0, x_1 \right\rangle w_2 - \frac{1}{2} w_2^T E\left\langle r_2 \,|\, x_0, x_1 \right\rangle \end{aligned}$$

To obtain the optimal decision of the second stage, we do the first order differential equal to zero on it.

$$2Var\left\langle r_2 \,|\, x_0, x_1 \right\rangle w_2 - \frac{1}{2} E\left\langle r_2 \,|\, x_0, x_1 \right\rangle = 0$$

$$w_2^* = \frac{1}{4}\left(Var\left\langle r_2 \,|\, x_0, x_1 \right\rangle\right)^{-1} E\left\langle r_2 \,|\, x_0, x_1 \right\rangle$$

According to Eq.8 and Eq.9 in section 3.1, the derived results as follows:

$$E\langle \mathrm{r}_2 \,|\, x_0, x_1 \rangle = E\langle \beta + \Pi_1 x_1 + u_2 \,|\, x_0, x_1 \rangle = \beta + \Pi_1 x_1$$

$$Var\langle \mathrm{r}_2 \,|\, x_0, x_1 \rangle = Var\langle \beta + \Pi_1 x_1 + u_2 \,|\, x_0, x_1 \rangle = \Sigma^*$$

The derived result is:

$$w_2^* = \frac{1}{4}(\gamma + \Phi x_1) \qquad \text{where } \gamma = (\Sigma^*)^{-1}\beta \text{ ; } \Phi = (\Sigma^*)^{-1}\Pi_1$$

Give a factor a_1 then becomes:

$$a_1 w_2^* = \frac{1}{4} a_1 (\gamma + \Phi x_1)$$

Next, take the optimal weight of the second phase, which multiplies an undetermined coefficient a_1, into the optimal objective function of two-stage method as shown below:

$$\begin{aligned} f_T(V) &= Var\langle V^T Z \,|\, x_0 \rangle - \frac{\lambda_{v2}}{2}\left[E\langle V^T Z \,|\, x_0 \rangle - \mu_O\right] \\ &= V^T G^* V - \frac{\lambda_{v2}}{2}\left[V^T F^* - \mu_O\right] \end{aligned}$$

The vectors and matrices represent respectively:

$$V = \begin{bmatrix} w_1 \\ a_1 \end{bmatrix} \;;\quad Z = \begin{bmatrix} r_1 \\ w_2^{*T} r_2 \end{bmatrix} \;;\quad F^* = E\langle Z \,|\, x_0 \rangle = \begin{bmatrix} E\langle r_1 \,|\, x_0 \rangle \\ E\left\langle w_2^{*T} r_2 \,\middle|\, x_0 \right\rangle \end{bmatrix} \;;$$

$$G^* = Var\langle Z \,|\, x_0 \rangle = \begin{bmatrix} Var\langle r_1 \,|\, x_0 \rangle & Cov\left\langle r_1, w_2^{*T} r_2 \,\middle|\, x_0 \right\rangle \\ Cov\left\langle w_2^{*T} r_2, r_1 \,\middle|\, x_0 \right\rangle & Var\left\langle w_2^{*T} r_2 \,\middle|\, x_0 \right\rangle \end{bmatrix}$$

To optimize the objective function, we do the first order differential equal to zero on it.

$$2G^* V - \frac{\lambda_{v2}}{2} F^* = 0$$

$$V^* = \frac{\lambda_{v2}}{4} G^{*-1} F^*$$

According to Eq.8 and Eq.9 in section 3.1, the derived results are as follows:

$$E(r_1) = E(\beta + \Pi_1 x_0 + u_1) = \beta + \Pi_1 x_0$$

$$E\left(w_2^{*T} r_2\right) = \frac{1}{4}\ell \quad \text{where} \quad \ell = L^T \Pi \beta + tr(\varphi)\ ; \quad L = (\gamma + \Phi\beta)\ ; \quad \varphi = \Phi^{\mathrm{T}} \Pi_1 \Sigma^*$$

$$Var\langle r_1 \,|\, x_0 \rangle = Var\langle \beta + \Pi_1 x_0 + u_1 \,|\, x_0 \rangle = \Sigma^*$$

$$Var\left\langle w_2^{*T} r_2 \,\middle|\, x_0 \right\rangle = \frac{1}{16}k \quad \text{where} \quad D = \Pi_1 \Sigma^* \Pi_1^T + \Sigma^*\ ; \quad \varphi = \Phi^{\mathrm{T}} \Pi_1 \Sigma^*\ ;$$

$$k = \gamma^T D \gamma + \left(\gamma^T + L^{\mathrm{T}}\right) D \Phi \beta + 2tr\left(\varphi^2\right) + tr(\varphi) + \left(\beta^T \Pi^T \Phi + 2L^T \Pi_1\right) \Sigma^* \Phi^{\mathrm{T}} \Pi \beta$$

$$Cov\left\langle r_1, w_2^{*T} r_2 \,\middle|\, x_0 \right\rangle = \frac{1}{4}N \qquad \text{where} \quad N = \Sigma^* \left(\Pi_1^T \gamma + \Pi_1^T \Phi \beta + \Phi^{\mathrm{T}} \Pi \beta\right)$$

Use the formula, $E\left\langle V^{*T} Z \,\middle|\, x_0 \right\rangle = \mu_O$, to solve λ_{2v}.

$$\mu_O = E\left\langle V^{*T} Z \,\middle|\, x_0 \right\rangle = \frac{\lambda_{v2}}{4} F^{*T} G^{*-1} F^*$$

$$\lambda_{v2} = \frac{4\mu_O}{F^{*T} G^{*-1} F^*}$$

The vector V_λ^* that represents the combination of the optimal weight of the first stage and an optimal factor of the second stage is:

$$V_\lambda^* = \frac{\mu_O}{F^{*T} G^{*-1} F^*} \cdot G^{*-1} F^*$$

The variance of the two-stage method as follows:

$$Var\left\langle V_\lambda^{*T} Z \,\middle|\, x_0 \right\rangle = \left(\frac{\mu_O}{F^{*T} G^{*-1} F^*} \right)^2 \cdot F^{*T} G^{*-1} G^* G^{*-1} F^*$$

$$= \frac{\mu_O^2}{F^{*T} G^{*-1} F^*}$$

Appendix2.4 · $\Lambda = 0; \beta = 0$

When x_0, x_1 are knew, the weight of the second stage is $w_2^* = \Phi x_1$ where $\Phi = (\Sigma^*)^{-1}\Pi_1$. While given a coefficient, a_1, it becomes $a_1 w_2^* = a_1 \Phi x_1$. Then apply the optimal weight of the second phase into the optimal objective function of the two-stage method as shown below:

$$\begin{aligned} f_T(V) &= E\left\langle V^T Z \,\middle|\, x_0 \right\rangle - \frac{\lambda_8}{2}\left[Var\left\langle V^T Z \,\middle|\, x_0 \right\rangle - \sigma_0^2\right] \\ &= V^T F_0 - \frac{\lambda_8}{2}\left[V^T G_0 V - \sigma_0^2\right] \end{aligned}$$

The vectors and matrices represent respectively:

$$V = \begin{bmatrix} w_1 \\ a_1 \end{bmatrix};\quad Z = \begin{bmatrix} r_1 \\ w_2^{*T} r_2 \end{bmatrix};\quad F_0 = E\langle Z \mid x_0 \rangle = \begin{bmatrix} E\langle r_1 \mid x_0 \rangle \\ E\left\langle w_2^{*T} r_2 \,\middle|\, x_0 \right\rangle \end{bmatrix};$$

$$G_0 = Var\langle Z \mid x_0 \rangle = \begin{bmatrix} Var\langle r_1 \mid x_0 \rangle & Cov\left\langle r_1, w_2^{*T} r_2 \,\middle|\, x_0 \right\rangle \\ Cov\left\langle w_2^{*T} r_2, r_1 \,\middle|\, x_0 \right\rangle & Var\left\langle w_2^{*T} r_2 \,\middle|\, x_0 \right\rangle \end{bmatrix}$$

After taking the first order differential equal to zero, we get:

$$V^* = \frac{1}{\lambda_8} G_0^{-1} F_0$$

Here the vector F_0 and the matrix G_0 are the same as described earlier in the paper.

Then solve for λ_8 and get the expected return as follows:

$$E\left\langle V_\lambda^{*T} Z \,\middle|\, x_0 \right\rangle = \sigma_O \sqrt{F_0^T G_0^{-1} F_0}$$